LEMOINE, SAMUEL J., M.S. The Development and Implementation of a Novel Bioisosteric Software Program for the Development of New Pharmaceuticals. (2014) Directed by Dr. J. Phillip Bowen. 94 pp.

This thesis is focused on the development of some important informatics methods that will assist drug discovery scientists in their pursuit of novel molecular structures, and it is representative of a novel information-based approach to drug design. This work utilized *in silico* based strategies, and did not involve standard computational chemistry energy-based calculations (e.g., molecular mechanics, quantum mechanics, or molecular dynamics simulations). The cheminformatics approach described herein uses empirically derived "rules", based on historically successful lead generation transformations where an original molecular structure served as the basis for a new one with enhanced biological activity. In some cases these substitutions, replacements, and/or new scaffold ring systems transformations may have been overlooked by pharmaceutical scientists involved in the systematic trial-and-error approaches commonly found in drug discovery. The novel "rule"-based transformations described serve as a basis for drug discovery and should serve as a resource for scientists wanting to explore chemical space based on previous examples of molecular modifications.

THE DEVELOPMENT AND IMPLEMENTATION OF A NOVEL BIOISOSTERIC SOFTWARE PROGRAM FOR THE DEVELOPMENT OF NEW PHARMACEUTICALS

by

Samuel J. Lemoine

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2014

Approved by

_____
Committee Chair

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGEMENTS

The author would like to thank all of those that have provided assistance during the years that this thesis has been researched. While the list of advisors and helpers is too extensive to thank personally, special gratitude goes to Dr. J. Phillip Bowen and the members of the Bowen Research Group (now located at Mercer University) who have graciously given of their time and expertise. Also, special thanks to the many members of the faculty at the University of North Carolina at Greensboro that have assisted with their time and energies, including Dr. Jason Reddick, Dr. Bruce Banks, Dr. Mitch Croatt, Dr. Ed Shamshoum, Dr. Greg Raner, and Dr. Jerry Walsh.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER I

BIOISOSTERES AND SUBSTITUTIONS

The drug discovery and approval process are extremely costly and time consuming. The expenses associated with FDA approval are approaching 1.5 billion dollars, which has impacted the way in which pharmaceutical companies approach problems. New lead molecular entities are run through a gauntlet of *in vitro*, *in vivo*, and *in silico* tests to determine their suitability for human use prior to undertaking any clinical trials. Currently most of the drug failures are related to poor pharmacokinetic or toxicity problems, which have been discovered in the late stages of clinical trials. It would be much more preferable from an economic standpoint if drug candidates doomed to fail ultimately in clinical trials can be identified in the earlier stages of drug discovery. Rarely does a drug come to market without extensive modifications having been made to the original lead molecular structure. More often than not, the drug has undergone a series of systematic modifications designed to improve any known pharmacologic and/or toxic deficiencies that have been detected in laboratory and animal model studies. Sometimes these modifications are minor, and other times, there may be so many changes that a completely new compound has emerged that seemingly has little or no similarity to the original lead.

For those lead compounds that do show promise, experimentalists will often make small changes to improve the necessary list of characteristics that are necessary for improved efficacy and pharmacokinetic patterns. These small incremental changes and subsequent biological testing are the basis for structure-activity relationships, which provide drug discovery scientists with information about what molecular features enhance or decrease biological activity. Most commonly, medicinal chemists focus on systematically modifying one part of a potential drug candidate, while leaving the rest of the structure unmodified. Such approaches have yielded commercially successful drug therapies. Perhaps the most highly acclaimed example of such systematic modifications based on physical organic chemistry is cimetidine, which is used for the treatment of stomach ulcers and related gastrointestinal disorders. Cimetidine was the first billion dollar selling medicine. It is not unusual for pharmaceutical laboratories to replace hundreds of substituents in an attempt to find the most effective structure. The basis for many of these changes include historical precedent, ease of synthesis, information gleaned from structural studies, and physicochemical reasons. To find the right substituent, researchers have often invoked the idea of bioisosteres.

Bioisosteres are substituents that are similar in their chemical and/or physical properties, and they achieve comparable biological functionality. Bioisosteres are classified into two categories: *classical* and *non-classical* bioisosteres. Classical bioisosteres are frequently similar in size (although it is not absolutely necessary in all cases), so they may be added to a lead structure without in general causing any new,

unfavorable internal steric interactions and more importantly without interfering with drug-receptor binding. An example of a bioisosteric replacement would be uracil and 5-fluorouracil. The two molecules are identical with the exception of a hydrogen atom that has been replaced by a fluorine atom at the double bonded carbon of the ring structure. The atomic radius of the fluorine atom is similar in size to a hydrogen atom, but there is a completely different electronic environment created. Such a substitution, fluorine for hydrogen, may serve to provide a new favorable electrostatic interaction with the receptor (a pharmacodynamics effect), or it may serve to provide a metabolic blocker (a pharmacokinetic effect). In the specific case of 5-fluorouracil, the fluorine-hydrogen substitution results in a mechanism-based inactivator.

Clearly in the case of the substitution of a hydrogen with a fluorine the difference in the electronegativity has the potential for a significant impact. Although 5-fluorouracil is similar in shape to uracil, the chemistry and pharmacology of the two drugs provide a different enzymatic outcome. In the case of uracil, the natural mechanistic pathway associated with thymidine synthase is removal of a proton. When the proton has been replaced with a fluorine, a suicide substrate inhibitor has been created because the fluorine, the most electronegative element, cannot leave as a positively charged ion (as is the case with a hydrogen in that position of the ring) which is a requirement for the biochemical transformation. This simple substitution that produces 5-fluorouracil generates a drug that is able to inhibit RNA replication enzymes, which helps prevent the growth of cancerous cells[1].

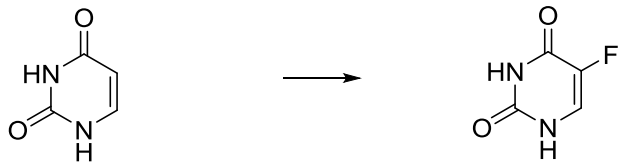

Figure 1. A side-by-side comparison of the natural substrate (uracil, 1) for thymidate synthase and the anticancer drug 5-fluorouracil.

It should be noted, that the programming language used for the development of the "rule"-based transformations outlined in this thesis automatically replaces Kekule structures with ring structures. For example, the single Kekule strcuture of benzene on the left in Figure 2 would be displayed only as the structure on the right. This is even true for molecular systems that may not be classically considered as aromatic, which is the case in Figure 1.

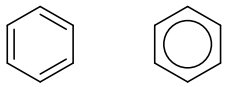

Figure 2. A comparison of aromatic structures of benzene.

Another example of a classical bioisostere is found in comparing procainamide and procaine. When comparing the structures of procaine and procainamide, it is clear that the only substitution between these two compounds is a nitrogen atom replacement of an oxygen atom (i.e., an ester functional group has been replaced by an amide). It should be noted that the arrow in Figure 3 and throughout this thesis are not implying necessarily that there is a single chemical transformation to convert one structure into another. Rather, the arrows displayed are depicting an *in silico* transformation that yields

another potential lead structure. Therefore, these reaction arrows are not connecting reactants to products, but original structures to new structures that should be considered as potential leads. While this substitution in Figure 3 may not seem significant, there is a dramatic difference in reactivity due to electronic differences. Esters hydrolysis occurs more readily than amide hydrolysis. The change from ester to amide has resulted in the creation of a drug that has been used in the treatment of cardiac arrhythmias for over sixty years.[2] Again, it should be noted that an actual chemical transformation converting an ester to an amide is not what is meant in Figure 3. What is being shown is the comparison of two similar molecular structures which typically must be prepared synthetically by two different routes.

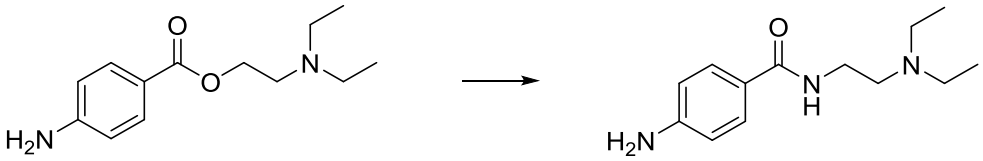

Figure 3. Procaine to procainamide transformation.

Classical bioisosteres are generally ring equivalent structures, covalent atoms, or functional groups. Non-classical bioisosteres do not follow the same patterns as classical bioisosteres. While they could have the same number of atoms, their abilities to substitute are based upon characteristics such as electronic properties, spatial arrangements, physicochemical properties, or functions that are critical in biological activites[3]. The biological activity is the ultimate test for non-classical bioisosteres. A tetrazole moiety would be an ideal non-classical bioisosteric replacement for a carboxylate.[4]While the

structures of these two groups are vastly dissimilar from an atomic functional group perspective, their overall steric and electrostatic properties are quite similar from the perspective of drug-receptor interactions. This similarity for drug-receptor interactions leads to equivalent or enhanced biological activity. Thus, the tetrazole moiety often can be substituted for the carboxylate functional group. This isosteric replacement has been used effectively in the angiotensin building antagonists (e.g. losartan and irbesartan).

CHAPTER II

SMILES AND SMIRKS: THE DRUG GURU PROJECT

Historically, beginning with the work of Paul Ehrlich, pharmaceutical scientists have used a systematic trial-and-error approach in new discoveries. Before the advent of recent technologies, many of these scientists worked independently, or in small groups, in developing new compounds and treatments, with little or no knowledge of the work of their contemporaries. They relied upon print publications and word of mouth to keep informed. Many chemists have spent a great deal of time working with promising lead compounds that have not resulted in viable drug candidates because potential drug modifications were ignored or overlooked. In many other cases, a new discovery could have been made with a slight modification of the drug structure.

Drug discovery demands a collaborative environment traditionally within a proprietary corporate research setting. As computer technology advanced in the latter part of the twentieth century, it has become much easier for researchers to work in different laboratories in different parts of the world to collaborate and share ideas. Chemists can share documents, experimental designs, and results within minutes of completion. Previously, other researchers would only be able to see these same results after reading them in a publication. But, even technology, however, has its limits. While chemists can share information, it would only be shared if the right people knew who

to speak with and where similar research was being conducted. Near the end of the twentieth century, scientists began to seek a way to make this happen.

In the late 1980's, chemist David Weininger and his associates designed a computer program called SMILES (Simplified Molecular Input Line Entry System). SMILES is a software program that utilizes a series of codes that replace letters and symbols for the elements and bonds. The coding is based upon the American Standard Code for Information Interchange (ASCII). Manipulation of this coding allows for accurate representation of elements, bonding, aromaticity, branching, stereochemistry, and isotopes. By using Structure Diagram Generation algorithms, two dimensional figures can be created that show graphical representations of the compounds shown.

The Drug Guru project was developed at Abbott Laboratories (now AbbVie) to assist medicinal chemists with the development of novel lead structures. This clever and useful approach utilized SMILES coding. Dr. Kent Stewart (AbbVie) developed a computer program in the mid 2000's that was designed to provide Abbott researchers with an opportunity to benefit directly from the efforts of other scientists, even without an active collaboration.[6] The central role was to develop a set of rules that could generate a series of novel lead structures based on the success stories of other drug discovery projects. Experienced medicinal chemists know many of the transformations that can be incorporated into a computer code "rules of thumb" using the SMILES language (e.g., the conversion of an ester to an amide, the conversion of a hydroxy to methoxy group, etc.)[7] The Drug Guru program takes these transformations and converts them to a web-based

application that utilizes computer codes to represent each transformation. These "rules" are represented as SMILES code (and SMIRKS, which is based upon the SMILES platform, although the name "SMIRKS" is used only because of its similarity to "SMILES" and isn't actually an acronym).[8] When a "rule" has been deemed generally useful, it is entered into the program and stored for future applications for lead structures possessing the potential functional group or molecular structure transformations. A "rule" does not guarantee that the new structure will have biological activity or is even appropriate for the specific case, but it is a possibility that should be considered. Drug discovery scientists are looking routinely to develop new biologically active compounds based on a current lead. An automated system for structure generation built on a database of transformations is an excellent starting point. This is a tremendous resource for a medicinal chemist with limited experience. Even experienced medicinal chemists, however, should be able to benefit from an automated program of this type inasmuch as they may overlook possible families of compounds. Moreover, such an approach may be extremely beneficial in circumventing patent protection issues for existing drugs.

In 2006 there were 186 general rules in the Drug Guru system. Of these rules, 133 were functional group transformations (e.g., 18 amide transformations), and the remaining 53 were molecular framework modifications (e.g., 14 ring break transformations). The typical input search at that time yielded between 50-150 output structures. The size of the input structure was important. Smaller organic molecules will yield fewer output structures in general, while larger more complex molecules will

provide many more options for output structures. Obviously, as the size of the lead molecule increases, there are more possibilities for the application of the transformation rules. Admittedly, experienced medicinal chemists already know many, if not all, of the fundamental rules. As mentioned above, Drug Guru is more of a reminder to these chemists particularly as more rules are added to the database that represent less known transformations. Previously researchers at Abbott Laboratories have identified a large number of these transformations since the development of the software in 2006. The current version (as of 2011) lists 530 transformations.[7] Despite this progress, there are many thousands of transformations that have yet to be identified and put into the SMILES and SMIRKS coding for Drug Guru.

CHAPTER III

LIMITATIONS

As with any other scientific method, there are limitations to this type of informatics approach for lead development. One of the major criticisms is that it not an artificial intelligence program. While a transformation might be perfectly suitable in one research project, it may not apply in another. The transformations do not recognize this problem. The software will apply whatever rules have been programmed into the system and generate new structures accordingly, regardless of the target system being explored. While it will recognize problems with bonding, aromaticity, and charges, it will not identify specific issues with impracticality, conformational preferences, and energy differences. So, it cannot be a substitute for the "common sense" of the researcher nor can this approach be considered a structure-based drug design approach. The chemist must exercise caution when reviewing the output structures based on this rule-based system. In developing the cheminformatics rules, one has to be careful to check the results for unintended consequences, which are molecular structures that are chemically incorrect (i.e., nonsensical structures). This can easily happen for generalized rules even if the rules have been checked thoroughly.

Some medicinal chemistry researchers have complained that there are too many new structures produced, while others have lamented that not enough diverse possibilities

have been generated. The chemists who have complained about too many potential leads have voiced their concerns that the sheer volume of possibilities is overwhelming. In some "rules of thumb" cases, there are indeed literally hundreds of known substitutions. There is also no ranking scheme based on some penalty function, so a chemist may have to sift through dozens of suggestions to determine which one is the best by inspection. Algorithms have yet to be developed in this context that will allow the suggestions to be ranked, although future plans in the Bowen group are to develop such penalty functions.

The chemists who have complained about a lack of quantity, generally have worked with compounds that have very few known substitutions or have had little research published on the subject. It can be quite frustrating to enter the information for the compound and get only one or two output structures. In the worst case scenarios, there are no substitutions offered. The value of the system to these researchers is limited.

In the case of Drug Guru another criticism has emerged about the user friendliness of the software. While the SMILES and SMIRKS codes make sense, the software itself may be cumbersome and difficult to maneuver. If the chemist understands the SMILES and SMIRKS codes, but cannot navigate the software platform to use these codes, the program becomes useless. Countless error messages and technical jargon problems can convince many to avoid the use of this method. Many chemists would rather fight their battles in the laboratory where they have extensive knowledge, rather than fight with a computer. This means that any software designed must take into account not only the functionality required but the computer skills of the intended user –

the bench chemist involved in medicinal chemistry synthesis, who may or may not have the computer programming skills necessary to navigate a complex cheminformatics program.

CHAPTER IV

A NEW SOFTWARE PROGRAM

The Drug Guru code is proprietary, and a decision was made at Abbott not to release it to the public. A new software program utilizing a similar rule-based strategy is being developed by the Bowen research group in the Center for Drug Design at Mercer University College of Pharmacy. The premise of the software program is simple. As researchers design new compounds for drug design, they frequently encounter stumbling blocks. These stumbling blocks may include side effects, insufficient or overabundant inhibition, binding specificity, reaction specificity, rate acceleration, lipophilicity effects, steric effects, enzyme inactivation, and electrostatic binding among the many problems that may be encountered. In testing, a new compound might show promising biological activity, but needs one or more minor modifications to be successful. The new software program will allow chemists to have an easier time with these modifications by suggesting proven substitutions. Presently there is no ranking scheme to prioritize the ranking structures, and these substitutions are not target specific. As the software continues to be developed, many of these concerns presumably will be addressed.

As mentioned previously, the concept of making substitutions is not new. A replacement of a small molecule with another small molecule has been the way drug discovery has historically worked. In the past, however, if a researcher had an issue with

a particular compound, the only way to decide on a good substitution was either trial-and-error hunches, or to rely on the expertise of the medicinal chemist who could apply knowledge-based decisions from experiences, precedents, and/or algorithms (e.g., Topless schemes). Ideally, if the knowledge of many experienced medicinal chemists could be distilled into a computer algorithm, then there would be a wealth of information available for anyone who used the software. This novel informatics approach provides the chemist with an established database of substitutions that have been successful in drug discovery case histories. So while a medicinal chemist may have an exceptional command of the drug discovery literature, a well-developed set of "rule-based' transformations will aid in the development of potential novel leads. If a researcher wanted to modify a particular part of a compound with new functional groups or make ring system alterations, he/she could simply input their compound into the new program and a list of suggested substitutions would be generated. For the veteran medicinal, some of these substitutions would be obvious, while others might be more obscure or simply overlooked. After the output list of structures is generated, they can be ranked according to other criteria (e.g., logP, Lipinski's Rules of Five, etc.). A review of the list of generated molecular structures should be undertaken to see if any make sense from a pharmacologic standpoint. No information about the synthesis of potential compounds would be provided because the set of transformation rules are not designed for this purpose. The transformations simply generate new molecular structures based on historical precedent and are not target specific. Obviously, not every substitution will

work in every situation. The database of rules has no way of "knowing" exactly what the researcher's intent might be. It merely provides a list of possibilities based on previous successful examples. Nevertheless, with that said, by providing these *in silico* generated possibilities, there is the potential to provide new leads for a project as well as reduce the time and expenses involved. One overlooked compound could be the difference between a successful drug candidate and the project being abandoned. A number of the output structures could be rejected for a variety of scientific and/or economic reasons. The success of a drug discovery project could rest on an obscure substitution. More often than not, the bioisosteres that have been studied are non-classical in nature. As research continues, the classical bioisosteres are mostly the 'rules of thumb', and the non-classical bioisosteres are the basis of new discovery.

One of the goals of this research project is to have 50-150 output structures for each input structure. For Drug Guru this number of output structures was achieved with only 530 rules in place. In just a short period of time, the Bowen research group has created a large set of rules that are generally applicable in drug discovery. Since we do not have access to the propriety code at AbbVie, assuming that our new rules are not already present in Drug Guru, this would increase that number of output structures by nearly 30%. With realistic projections, it would not be unreasonable within a short time that the Bowen research group would be able to double the 530 rules that were in place in 2011. With continued development, it is not unreasonable to expect that the rules could number in the thousands within a few short years. Such advances would make the new

program an increasingly valuable resource for the bench chemist, who would benefit the most from an easy to use program.

How is the database of transformations created? Like essentially everything else, the database is constructed on foundational principles. There are a few tried and true "rules of thumb" that serve as the building blocks of the database (some of these were discussed in the "Bioisosteres and Substitutions" section). The majority of the database is based on a careful evaluation of successful case studies found in the scientific literature. Each year, hundreds of research articles are published detailing successes and failures in laboratories from around the world. In a number of these publications and/or disclosures, substitutions have been made to improve pharmacokinetic and/or pharmacodynamics properties of a particular drug. For years, the Bowen research group has been meticulously screening data presented in publications and patents to identify the most general lead to drug transformations. When a substitution or modification has been made, the specifics of the transformation have been incorporated into SMILES and SMIRKS codes within the framework of the software program.

For the work of this thesis, the critical evaluation of the literature has been the most time consuming respect of the research. Typically, a monthly or bimonthly research journal may have anywhere between 20 and 50 research articles. While the abstracts can provide some basic information, they frequently do not go into depth about the substitutions that were made. In order to understand the physicochemical properties and biological activities that have been reported, each article must be analyzed. Such careful

reviews require substantial time investments, with no guarantee of finding anything that fits the criterion of the software. Additionally, when a suitable substitution has been identified, the SMILES and SMIRKS codes must be produced. There are some commercially available software programs that do assist in this endeavor, but they have limitations. In most cases, the code is constructed through trial-and-error methods with rigorous quality control checking. Just as a single bond or atom can wreak havoc on an otherwise stable molecule, a single misplaced symbol can mean success or failure in coding. Some codes are put together in literally seconds. In other cases, it takes several hours to get just the right framework to make the SMILES and SMIRKS codes meaningful for the desired transformation in a variety of chemical environments. The biggest challenge is making sure that there are no unintended consequences where modifications are made that lead to unrealistic structures. Chemistry drawing programs offer SMILES codes for structures input and export, but these codes are not always compatible with SMIRKS codes. Interestingly, the SMILES codes from one program are not always compatible with another, so this code porting issue creates additional frustrating obstacles. However, as research has continued over the past few years, the learning curve has been surmounted: the coding time has decreased, as practice and problem solving experiences have helped overcome prior frustrations and failures.

The final section of this thesis is devoted to the development of the cheminformatics rules that may be used in drug design. There are a large number of research articles referenced, with over two hundred SMILES and SMIRKS codes (along

with figures of each substituted portion). Approximately three thousand research articles were reviewed to select sufficiently general rules. It is not the intent of this thesis to present in detail the background and context for each transformation. The purpose of this thesis, however, is to provide some general transformation rules that have led to increased biological activities and are structurally interesting. These new rules should become part of the new database of transformations.

CHAPTER V

RESEARCH AND RESEARCH APPROACHES

Over the past few years, ongoing efforts in the Bowen lab have resulted in mastering a number of basic rules for Drug Guru and our in-house software. The transformations include 19 catechol transformations ("transformation" meaning in this case, a bioisosteric substitution), 17 carboxylate transformations, 7 aldehyde transformations, 10 amide transformations, 9 ester transformations, 8 ketone transformations, 3 thioether transformations, 3 thiourea transformations, and a number of others. These were important fundamental rules that will be important for the long term success of the program. A number of these substitutions are well known functional group replacement strategies (e.g., converting an ester to an amide or an inverse amide). Quite a few of the most fundamental transformations are detailed in Richard B. Silverman's textbook *The Organic Chemistry of Drug Design and Drug Action.*[9] These rules are the foundations of the system. Some selected examples of these well-known modifications are shown below. Each example shows the substitution of a hydroxyl group with some other functional group previously used in medicinal chemistry research. The boxes in the figures indicate the bonding points.

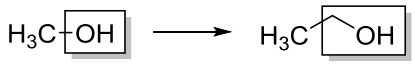

Figure 4. A hydroxyl group replacement with a methylhydroxyl group. The SMIRKS code for this substitution: CO>>CC

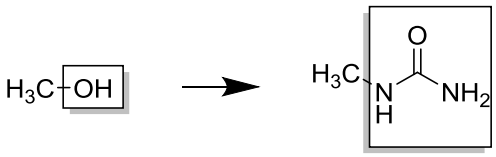

Figure 5. A hydroxyl group replacement with a 1-methylurea. The SMIRKS code for this transformation: CO>>CNC(=O)N.

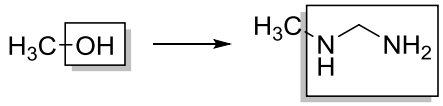

Figure 6. A hydroxyl group replacement with N-methylmethanediamine. The SMIRKS code for this transformation: CO>>CNCN.

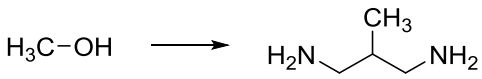

Figure 7. A hydroxyl group replacement with 2-methyl-1,3-propanediamine. The SMIRKS for this transformation: CO>>CC(CN)CN.

The following are some additional transformations showing “rules of thumb” from the

Silverman text:

Basic transformations (using SMIRKS codes)

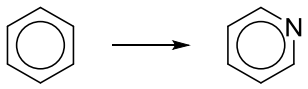

Figure 8. Benzene ring replacement with a pyridine ring. The SMIRKS code for this transformation: c1ccccc1>>c1ccccn1.

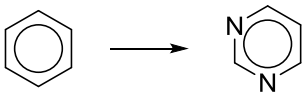

Figure 9. Benzene ring replacement with a pyrimidine ring. The SMIRKS code for this transformation: c1ccccc1>>c1ncncc1.

Figure 10. Benzene ring replacement with a ring-contracted polythiophene aromatic heterocycle. The SMIRKS code for this transformation: c1ccccc1>>c1sccc1.

Figure 11. Benzene ring replacement with a ring-contracted thiazole aromatic heterocycle. The SMIRKS code is: c1ccccc1>>c1cscn1.

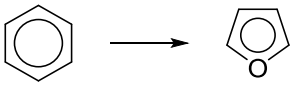

Figure 12. Benzene ring replacement with a ring-contracted furan aromatic heterocycle. The SMIRKS code is: c1ccccc1>>c1cocc1.

Figure 13. Benzene ring replacement with a ring-contracted isoxazole aromatic heterocycle. The SMIRKS code is: c1ccccc1>>c1conc1.

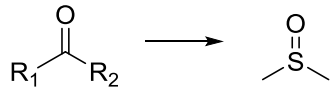

Figure 14. Carbonyl functional group replacement to dimethylsulfoxide group. The SMIRKS code is: C(=O)(*)(*)>>CS(=O)C.

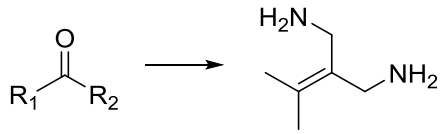

Figure 15. Carbonyl replacement with a 2-isopropylidene-1,3-propanediamine group. The SMIRKS code is: C(*)(*)=O>>CC(=C(CN)CN)C.

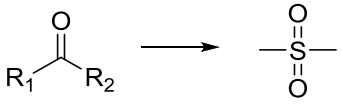

Figure 16. Carbonyl functional group replacement with a methylsulfonylmethane group. The SMIRKS code transformation: C(*)(*)=O>>CS(=O)(=O)C.

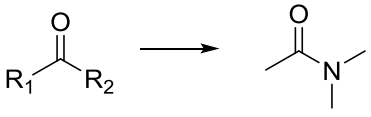

Figure 17. Carbonyl functional group replacement with a N,N-dimethylacetamide) aka "DMAc" group. The SMIRKS code transformation: C(=O)(*)(*)>>CC(=O)N(C)C.

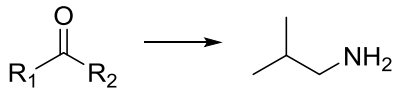

Figure 18. Carbonyl functional group replacement with an isobutylamine group. The SMIRKS code transformation: C(*)(*)=O>>CC(C)CN.

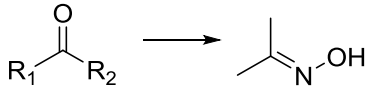

Figure 19. Carbonyl functional group replacement with a N-hydroxy-2-propanimine group. The SMIRKS code transformation: C(*)(*)=O>>CC(C)=NO.

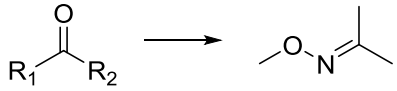

Figure 20. Carbonyl functional group replacement with a propan-2-one-O-methyl-oxime. The SMIRKS code transformation: C(*)(*)=O>>CON=C(C)C.

After several semesters of working on these fundamental transformations (some of which are listed above) the goal was to incorporate more advanced transformations using the latest cutting-edge transformations and/or structural modifications reported. The *in silico* transformations can be accomplished readily once the research has been published. The reason this strategy was adopted is at least twofold: (1) Transformations outlined in the Silverman text are essential to a rule-based transformation database because they are fundamental. (2) The development of these rules, typically while quite simple, provided the necessary background to learn the SMILES and SMIRKS language. Below is a summary of research.

The best method for developing new transformations is targeting recent substitutions and transformations that have been published in the leading chemistry and medicinal chemistry journals. Thus, an aggressive approach identifying useful transformations and writing them in a generalized way with SMILES and SMIRKS codes was the most logical plan of attack. To demonstrate our approach, an example of a recent transformation published in *Tetrahedron Letters* is shown.

Dr. Marc J. Adler and Dr. Steven W. Baldwin have been involved in acid-base protecting group research at Yale University and Duke University.[10] In 2009, in an effort to synthesize compounds involving 2,2-dimethylchromenes, a series of phenol based compounds was studied. The study showed a number of structural modifications in which 2,2-dimethylchromenes could be substituted for phenols. Considering the quantity of compounds that incorporate phenol rings, having a substitution which could change biological activity without radically altering the structure of the compound could be extremely valuable. The phenol ring (shown on the left of the Figure 21 below) was converted to the 2,2-dimethylchromene structure (shown on the right).

The transformation represents a potentially significant achievement. Having a transformation such as this within the Drug Guru and Bowen Research Group databases could be extremely helpful in a variety of areas. There is a simple SMILES and SMIRKS code that can be derived from this work. In order to maintain the integrity of the transformation, the benzene ring must be included in the code. On the product side, the new molecule has a ring formation that is attached to the benzene ring with an oxygen

atom enclosed and two additional carbons affixed. A SMILES code for this transformation would be c1c(O)cccc1>>CC1(C)Oc2ccccc2C=C1, while the SMIRKS code would be c1c(O)cccc1>>c1c(OC(C)(C)C=C2)c2ccc1. The two arrows (>>) represent the transformation, and they separate the left side and the right side of the equations. The lower case c's represent aromatic ring system carbon atoms. The (O) represents an oxygen atom that is attached to a carbon within that system. Note that the O that resides within the parentheses on the left hand side is also on the right hand side. However, the oxygen has now become part of a second, non-aromatic ring system. Those carbons are represented by upper case c(C). The two carbons that protrude from that second ring are represented by the (C)(C). When there is a double bond, it is represented by an equals sign (=). Inside the second ring structure, there is one double bonded pair of carbons. A triple bond would be represented by a '#', but none are shown in this particular transformation example. When there is nothing in between two letters, a single bond is assumed. Ring structures have a beginning point and an ending point. These are represented by the numbers (1,2). Therefore, in the SMILES shown, the first c has a 1 after it representing that this is the start of the aromatic ring, and after the sixth c, there is an additional 1, showing that this carbon is attached to the first carbon in the chain, thereby closing the ring. When this total transformation is viewed visually, it is easier to understand.

As shown below, the transformation can be seen in its entirety. The implications of this type of research can be significant. Is it possible that this substitution would work

in other reactions? By including this transformation in our databases, there is an option to exploit this information. The transformation is shown in Figure 21 below, with the substituted area in blocks.

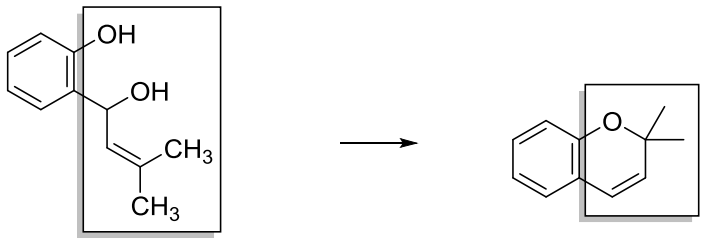

Figure 21. Compound 2-(1-hydroxy-3-methylbut-2-en-1-yl)phenol transformation to 2,2-dimethylchromene. The SMIRKS code transformation: c1c(O)c(C(O)C=C(C)C)ccc1>>c1c(OC(C)(C)C=C2)c2cc(O)c1.

Another example is taken from Dr. Christina Juli and her colleagues. They have been researching inhibitors to the Legionella MIP protein.[11]This protein is the primary cause for Legionnaire's Disease and Pontiac fever. It affects people with impaired respiratory systems. A number of substitutions were carried out with favorable changes in MIP activity. In the compounds shown in the fourth substitution, there is a $N(OH)_2$ group that is attached to a benzene ring coming off of a sulfur atom. The researchers substituted an amine in this location (as shown below in Figure 22). A number of promising MIP inhibitors were found (using this substitution and a few others), and the transformations that were performed could turn out to be useful in other studies.

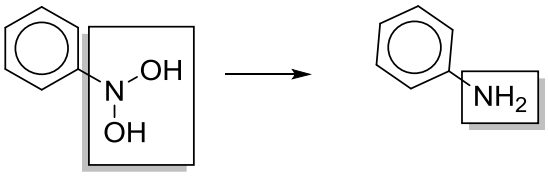

Figure 22. N,N-dihydroxyanaline to aniline. The SMIRKS code for this transformation: c1ccc(N(O)O)cc1>>c1ccc(N)cc1.

The conversion to aniline is similar to the SMILES/SMIRKS example shown previously. The benzene ring is signified by lower case c's representing an aromatic compound. Inside the parentheses are the attached groups. In the first side of the SMIRKS, the two oxygen atoms are represented. One is inside parentheses, and the other follows. This signifies that both oxygen atoms are attached to the nitrogen, but not to each other. On the right side of the equation, only the nitrogen is shown. This is because hydrogen atoms generally are not represented in the SMILES and SMIRKS codes as they are assumed to be there, much like chemistry line drawings.

CHAPTER VI

RESEARCH FIGURES AND DATA

The following reactions are grouped based upon the substituted compound. In some cases, the compound noted would fit in more than one category. All figures shown are SMILES based.

Benzanilide–Biphenyl Replacement: A Bioisosteric Approach to Quinoline Carboxamide-Type ABCG2 Modulators.[12]

UR-COP78 has been shown to be a potent ABCG2 modulator. Through replacement of the labile benzanilide with a biphenyl group, improved stability was achieved, while still maintaining potency and selectivity. The transformation keeps both aromatic rings intact (see Figure 23), but removes the nitrogen and oxygen atoms from the connecting chain, making the new molecule slightly smaller.

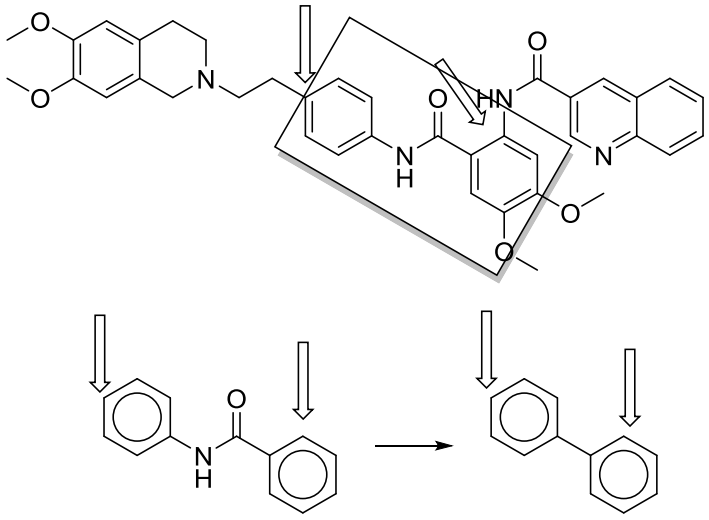

Figure 23. Benzanilide to biphenyl transformation (with parent structure and substituted area shown). The SMILES code transformation: O=C(Nc1ccccc1)c2ccccc2>>c1ccc(cc1)c2ccccc2.

SMIRKS:
[c:1]1[c:2][c:3][c:4]([c:5][c:6]1)C(=O)N[c:7]2[c:8][c:9][c:10][c:11][c:12]2>>[c:1]1[c:2][c:3][c:4]([c:5][c:6]1)[c:7]2[c:8][c:9][c:10][c:11][c:12]2

The bioisosteric similarity of the tetrazole and carboxylate anions: Clues from the topologies of the electrostatic potential and of the electron density.[13]

Researchers examined the tetrazole and carboxylate anions to compare how well one could replace the other as a bioisosteric substitution. They found that they were remarkably similar in their electrostatic potentials, despite the fact that they exhibit significant differences in molecular geometry and constitution. Clearly, one similarity is the ionizable hydrogen on both the carboxylic acid and the tetrazole. This type of

bioisosteric replacement is quite common in medicinal chemistry and is the basis for essentially all of the angiotensin II receptor antagonists.

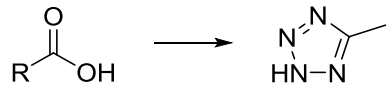

Figure 24. Carboxylic acid to 5-methyl tetrazole. The SMILES code transformation: CC(O)O>>CC1=NNN=N1.

SMIRKS:
[C:1][C:2](O)O>>[C:1][C:2]1=NNN=N1.

Benzothiazepine CGP37157 and its isosteric 2'-methyl analogue provide neuroprotection and block cell calcium entry.[14]

Benzothiazepine CGP37157 has recently been identified as a neuroprotector in mitochondrial cells. ITH12505 is an isosteric analogue of CGP37157, with a chlorine atom being replaced by a methyl group on the phenyl ring. Both the chlorine and methyl groups are hydrophilic. ITH12505 showed increased neuroprotective abilities. With continued research, it could be developed as an antioxidant in L-type voltage-dependent Ca^{2+} channels, where CGP37157 is currently being investigated. In the figure below, a chlorine atom has been removed from an aromatic ring and a methyl group has been added. Note that in the structure right structure below, the added methyl group is shown on the opposite side of the benzene ring, but this is the same position that the chlorine atom originally occupied if the benzene ring was rotated 180°.

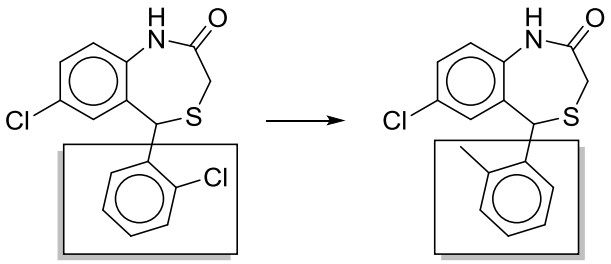

Figure 25. Compound 7-Chloro-5-(2-chlorophenyl)-1,5-dihydro-4,1-benzothiazepin-2(3H)-one to 7-chloro-1,5-dihydro-5-(2-methylphenyl)-4,1-benzothiazepin-2(3H)-one. The SMILES code transformation: Clc1ccccc1C3SCC(=O)Nc2c3cc(Cl)cc2>>Clc3cc1c(NC(=O) CSC1c2ccccc2C)cc3.

SMIRKS: Cl[c:1]1[c:2][c:3][c:4]cc1C3SCC(=O)Nc2c3cc(Cl)cc2>>Cl[c:1]3[c:2][c:3]1[c:4](NC(=O)CSC1c2ccccc2C)cc3

BACE-1 hydroxyethylamine inhibitors using novel edge-to-face interaction with Arg-296.[15]

Beta-secretase 1 (BACE1) is an enzyme in humans that is known to be involved in the advancement of Alzheimers Disease. A hydroxyethylamine template was used with substitutions to develop a potential inhibitor of BACE1. The interactions take place on the non-prime side of the molecule with Arg-296. In the transformation, a smaller ring structure has been replaced by a larger aromatic ring and a sulfonamide. When the substitution shown below was instituted, it increased inhibition in BACE-1$_{C50}$ by as much as thirty fold.

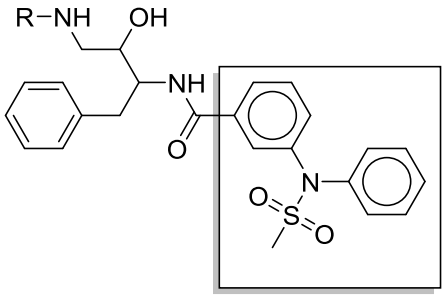

Figure 26. Diphenylamine BACE-1 inhibitor with substitution site indicated.

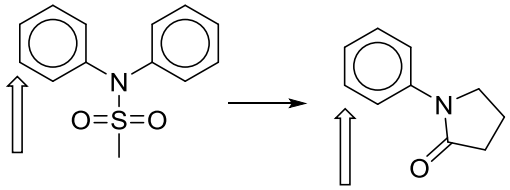

Figure 27. Compound N,N-diphenylmethanesulfonamide to 1-phenyl-2-pyrrolidinone. The SMILES code transformation:
CS(=O)(=O)N(c1ccccc1)c2ccccc2>>O=C2CCCN2c1ccccc1.

SMIRKS:
CS(=O)(=O)[N:1]([c:2]1[c:3][c:4][c:5][c:6][c:7]1)c2ccccc2>>O=C2CCC[N:1]2[c:2]1[c:3][c:4][c:5][c:6][c:7]1

Synthesis and antiproliferative evaluation of 5-*oxo* and 5-thio derivatives of 1,4-diaryl tetrazoles.[16]

L1210 is a leukemic cell, and SkBr3 is a cell line frequently associated with breast cancer. Researchers developed a series of compounds that utilized substitutions on an aromatic ring at the side of a 5-thio derivative. These compounds had early success in inhibition on both L1210 and SkBr3. Conceptually, the transformation of a tetrazole carbonyl moiety to a tetrazole thiocarbonyl moiety is shown in Figure 28.

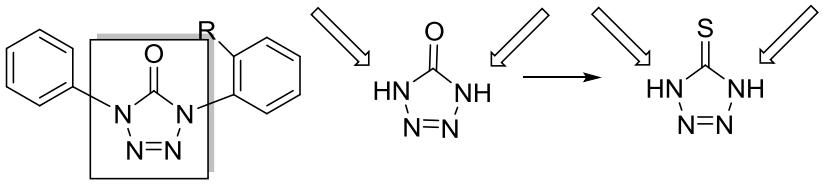

Figure 28. Compound 1,2-dihydro-5H-tetrazol-5-one to 1,2-dihydro-5H-tetrazole-5-thione transformation (with parent structure to the left). The SMILES code transformation: O=C1NN=NN1>>S=C1NN=NN1.

SMIRKS:
O=[C:1]1[N:2][N:3]=[N:4][N:5]1>>S=[C:1]1[N:2][N:3]=[N:4][N:5]1

Synthesis and biological evaluation of indolyl chalcones as antitumor agents.[17]

Indolyl alkaloids have recently been found to have significant effects on P388 leukemia cell cytotoxicity. Researchers tested a group of substituted indolyl chalcones for their anticancer activity against cell lines containing epithelial A-549, pancreatic carcinoma PaCa-2, and androgen-independent human prostatic adenocarcinoma PC-3 cancer cells. Two of the substitutions showed significant cytotoxicity against all three cells, but had especially strong results when working with the PaCa-2 cells. In the transformation, the ring structures remain constant, but the side chains attached have been changed.

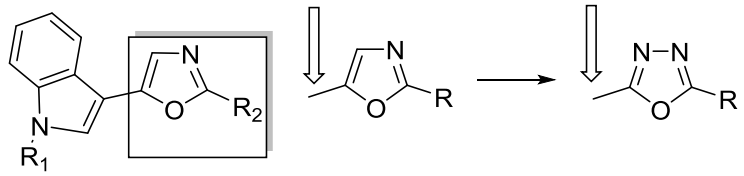

Figure 29. Parent structure with five membered ring structure substitution. The SMILES code transformation: CC1OC(*)=NC=1>>CC1OC(*)=NN=1

SMIRKS:
CC1OC(*)=NC=1>>CC1OC(*)=NN=1

Design of a series of bicyclic HIV-1 integrase inhibitors. Part 2: Azoles: Effective metal chelators.[18]

HIV integrase is a retrovirus enzyme that allows HIV to place its genetic material into the DNA of a cell. A series of isosteres was developed as promising HIV integrase inhibitors. The substitutions were based upon a ring structure with a sulfur-carbon exchange. In Figure 30 (shown below), the sulfone part of the [1,2,5]thiadiazolidine 1,1-dioxide ring is being replaced with a carbonyl and the CH_3 is being replaced with an isopropyl.

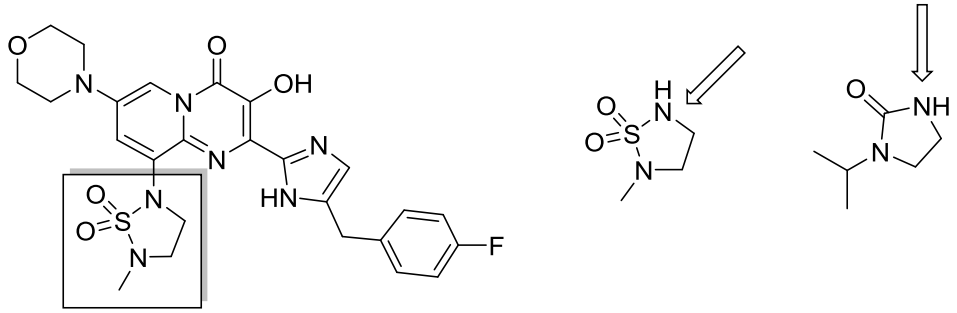

Figure 30. Parent structure with 2-Methyl-1,2,5-thiadiazolidine 1,1-dioxide to 1-Isopropyl-2-imidazolidinone substitution. The SMILES code transformation: CN1CCNS1(=O)=O>>CC(C)N1CCNC1=O.

SMIRKS:
[C:1][N:2]1[C:3][C:4][N:5]S1(=O)=O>>CC(C)[N:5]1[C:4][C:3][N:2][C:1]1=O

Trichoderins, novel aminolipopeptides from a marine sponge-derived *Trichoderma* sp., are active against dormant mycobacteria.[19]

Researchers isolated three aminolipopeptides from marine sponges in an effort to inhibit mycobacterial growth. The substitutions shown were determined to be the most potent in anti-mycobacterial activity. In Figure 31 (shown below), the transformation involves converting the unsaturated carbon-carbon double bond to the corresponding saturated system with the addition of a hydroxyl group bonded to the beta-carbon. Although the newly generated structure is the synthetic precursor of the enone structure and generally more unstable, there is precedent in having electronegative groups beta to carbonyl functional groups.

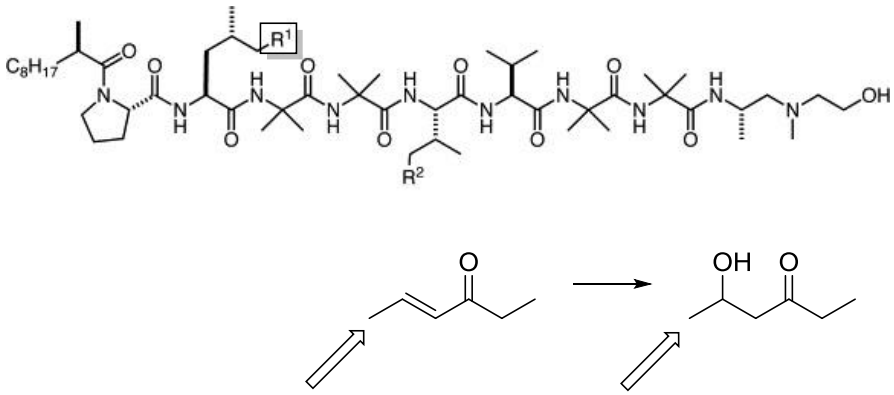

Figure 31. Parent compound with substitution site indicated followed by 5-Hydroxy-3-hexanone to (4E)-4-hexen-3-one transformation. The SMILES code transformation:C\C=C\C(=O)CC>>CCC(=O)CC(C)O

SMIRKS:
[C:1]\[C:2]=[C:3]\C(=O)CC>>[C:1][C:2][C:3](=O)CC(C)O

Synthesis, aerobic cytotoxicity, and radiosensitizing activity of novel 2,4-dinitrophenylamine tethered 5-fluorouracil and hydroxyurea.[20]

A series of substituted N-(2,4-dinitro-phenylamino) compounds was developed as HT-29 inhibitors. The hope for this is an increase in radiosensitizing abilities on cancer cell lines. The two compounds shown in Figure 32 were among a series of substituted compounds based on a 2,4-dinitrophenylamine scaffolding which increased cancer inhibition. The specific targets would be adenocarcinomas, which is a cancer of the epithelium. In this transformation, the methoxyurea was replaced with the fluorouracil-based compound, and it showed similar behavior in radiosensitizing. The transformation itself is a substitution of a smaller chain for a six membered aromatic ring.

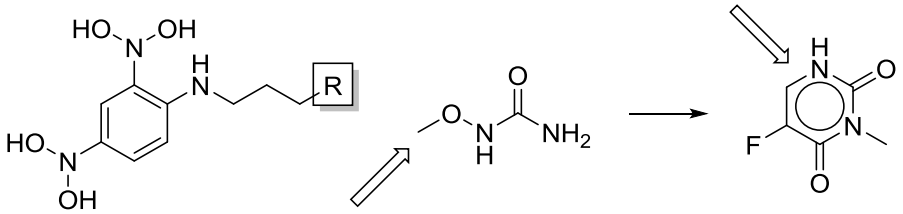

Figure 32. Parent structure with 1-methoxyurea to 5-Fluoro-3-methyl-2,4(1H,3H)-pyrimidinedione transformation. The SMILES code transformation:
CONC(=O)N>>Cn1c(=O)[nH]cc(F)c1=O.

SMIRKS:
CONC(=O)N>>Cn1c(=O)[nH]cc(F)c1=O

Benzodiazepine calcitonin gene-related peptide (CGRP) receptor antagonists: Optimization of the 4-substituted piperidine.[21]

In an effort to treat migraines, a series of substituted piperidines was developed as CGRP receptor antagonists. When combined with a dihydroquinazolinone and placed into a benzodiazepine core, these structures produced a stable, potent analog. As shown in Figure 33 below, the two ring structures are detached in the second compound.

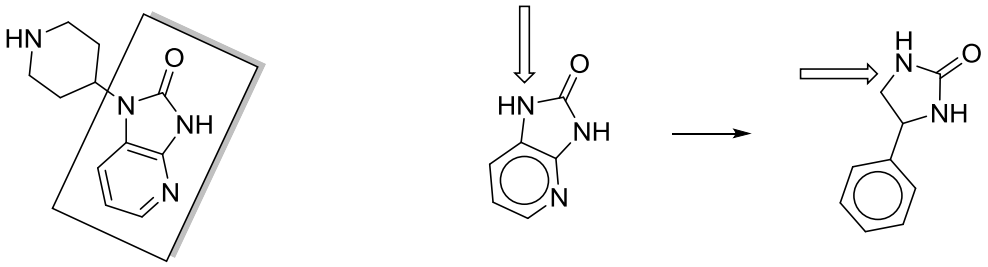

Figure 33. Parent structure with1,3-Dihydro-2H-imidazo[4,5-b]pyridin-2-one to 4-phenyl-1,3-dihydroimidazol-2-one substitution (with attachment points indicated).
The SMILES code transformation:
O=C1[NH]C2CCCNC2[NH]1>>O=C1[NH]CC([NH]1)C2CCCCC2

SMIRKS:
O=C1C2CCCNC2N1>>O=C1NC(=CN1)C2CCCCC2

Synthesis, Activity, and Pharmacophore Development for Isatin-β-thiosemicarbazones with Selective Activity toward Multidrug-Resistant Cells.[22]

β-Thiosemicarbazones have been found to be effective in the inhibition of P-glycoprotein. P-glycoprotein gives multi-drug resistance to cancer cells. The elimination of this protein should result in better anticancer treatment options. The modified compound shown below in Figure 34 has some promising inhibition results. The essential

transformations resulted in the removal of a carbonyl group and the opening of a ring. The benzene ring is part of the parent structure (the entire structures are shown in the figure).

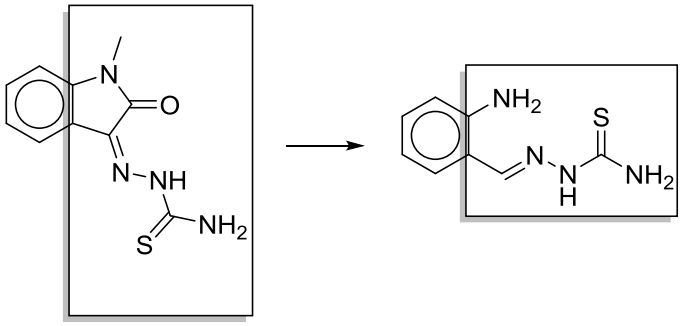

Figure 34. Methisazone ((2Z)-2-(1-methyl-2-oxo-1,2-dihydro-3H-indol-3-ylidene)hydrazinecarbothioamide) to (2E)-2-(2-aminobenzylidene) hydrazinecarbothioamide. The SMILES code transformation: S=C(N)N/N=C2/c1ccccc1N(C2=O)C>>S=C(N)N/N=C/c1ccccc1N.

SMIRKS:
S=C(N)N/N=C2/[c:1]1[c:2][c:3][c:4][c:5][c:6]1N(C2=O)C>>S=C(N)N/N=C/[c:1]1[c:2][c:3][c:4][c:5][c:6]1

New Analgesics Synthetically Derived from the Paracetamol Metabolite *N*-(4-Hydroxyphenyl)-(5*Z*,8*Z*,11*Z*,14*Z*)-icosatetra-5,8,11,14-enamide.[23]

N-(4-hydroxyphenyl)-(5Z,8Z,11Z,14Z)-icosatetra-5,8,11,14-enamide (AM404) is an inhibitor of endocannabinoid cellular uptake. By shortening the acyl chain, bioisosteric compounds were developed that had similar analgesic activities. The most significant transformation in this case was the removal of the hydroxyl group and the

addition of another heteroaromatic ring. Presumably the indazole nitrogen can supply a pair of electrons similar to the hydroxyl group.

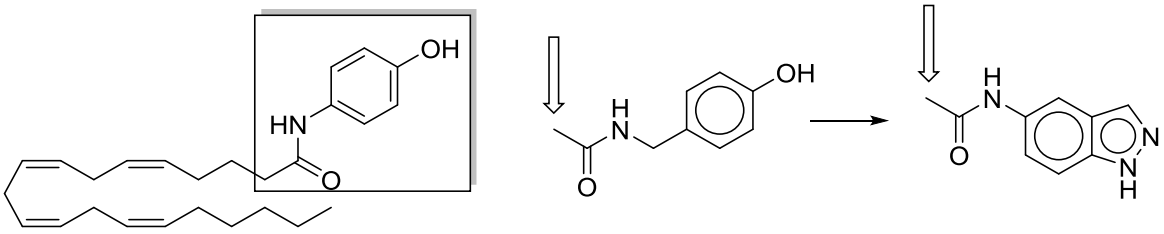

Figure 35. Compound AM404 with N-(4-hydroxybenzyl)acetamide to N-(1H-indazol-5-yl)acetamide substitution. The SMILES code transformation:
O=C(NCc1ccc(O)cc1)C>>CC(=O)Nc1cc2cnnc2cc

SMIRKS:
O=[C:1]([C:2])NCc1ccc(O)cc1>>[C:1][C:2](=O)NC1=CC2C=NNC2C=C1

Effect of Thioxopeptide Bonds on α-Helix Structure and Stability.[24]

Researchers have studied the effect of thioxoamides on protein structure and function. The thioxoamide bonds are similar in structure to amide bonds, but have different photophysical and hydrogen-bonding properties. Experimental results found that the substitution is tolerated, but there is some destabilization with the substitution. Given that the replacement of a carbonyl oxygen with a sulfur is an isosteric replacement, there are a number of situations in which a C=O could be substituted with a C=S.

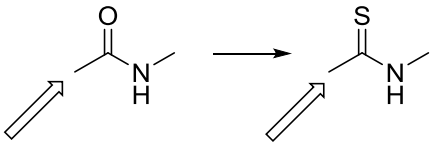

Figure 36. N-Methylacetamide to N-Methylethanethioamide substitution with attachment points to an α-Helix (not shown) structure as indicated. The SMILES code transformation: O=C(NC)C>>S=C(NC)C

SMIRKS:
O=[C:1]([N:2][C:3])[C:4]>>S=[C:1]([N:2][C:3])[C:4]

Cyclic amide bioisosterism: Strategic application to the design and synthesis of HCV NS5B polymerase inhibitors.[25]

HCV NS5B polymerase has been identified as being an important factor in the Hepatitis C virus. In an effort to create successful thiophene carboxylate inhibitors of HCV NS5B polymerase, bioisosteric modeling was employed to create compounds that were compatible with the original amide inhibitors. In Figure 37, a triazole was substituted for a non-aromatic compound.

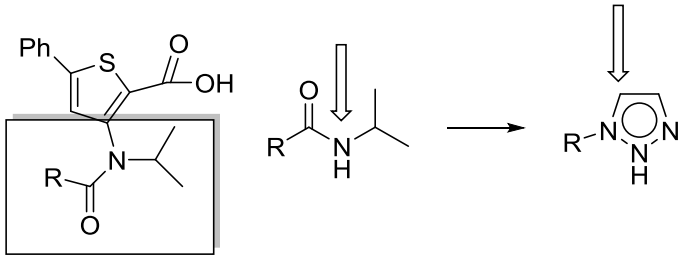

Figure 37. Parent compound with amide-based bioisosteres and attachment points shown. The SMILES code transformation: CC(C)NC(*)=O>>*n1ccnn1.

SMIRKS:
CC(C)NC(*)=O>>*n1ccnn1

Synthesis and biological evaluation of indolyl glyoxylamides as a new class of antileishmanial agents.[26]

Leishmania is a disease that is transmitted by insect bites. Through the bite wound, a parasite is released into the host's body. Eventually the parasite attacks vital organs. It has been proven to be very drug resistant. By synthesizing indolyl glyoxylamide derivatives, antileishmanial activity was increased. The substituted groups shown below (Figure 40) were attached at the bottom of a tricyclic ring structure as shown in Figure 39. The two molecules are very similar in size and shape, but the substitution of a bromine atom for an ethyl group gives it different properties.

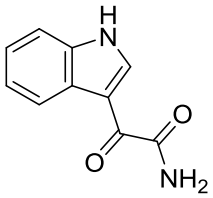

Figure 38. Indole-3-glyoxylamide is shown.

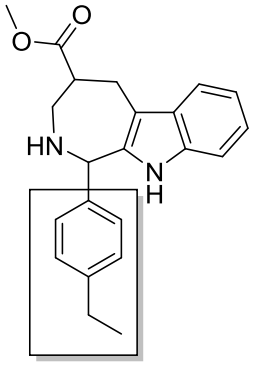

Figure 39. Connection point for the substituted ethylbenzen

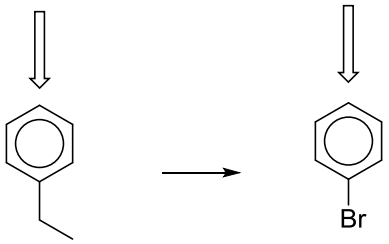

Figure 40. Ethylbenzene to bromobenzene transformation
The SMILES code transformation: CCc1ccccc1>>Brc1ccccc1 .

SMIRKS:
c1(CC)ccccc1>>Brc1ccccc1

1,2-Diamines as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1).[27]

N(1)-benzyl-N(1)-methylethane-1,2-diamine was used as a substitute for (S)-alanine benzylamide in the design of a CARM1 inhibitor. The simple substitution shown below attached to the diamine as listed, increased potency by greater than a factor of three with improved half-lives.

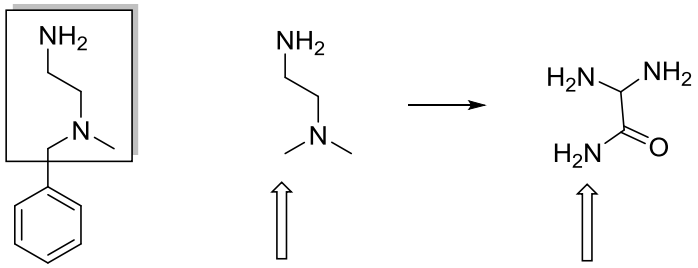

Figure 41. N(1)-benzyl-N(1)-methylethane-1,2-diamine on the left with the substitution location boxed. N-methylethane-1,2-diamine to 2,2-diaminoacetamide substitution shown with attachment points in arrows. The SMILES code transformation: CNCCN>>NC(N)C(N)=O.

SMIRKS:
NCCNC>>NC(N)C(N)=O

Heterocyclic acetamide and benzamide derivatives as potent and selective β_3 -adrenergic receptor agonists with improved rodent pharmacokinetic profiles.[28]

In an effort to treat symptoms involved in an overactive bladder, a series of adrenergic receptor agonists was investigated. By isosteric substitution, selectivity within these agonists was significantly improved. In Figure 42, a number of substitutions have been made (including the introduction of a sulfur atom), but the size and shape of the molecules are very similar.

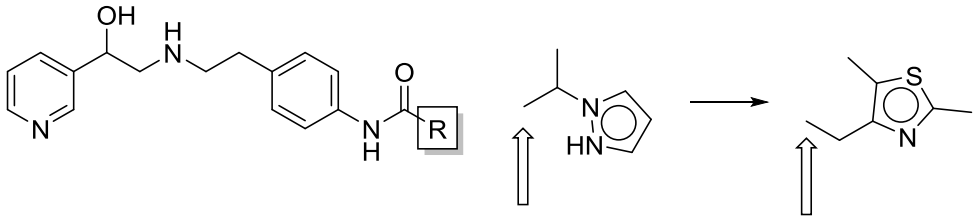

Figure 42. Parent structure with 1-(propan-2-yl)-1*H*-pyrazole to 2,5-dimethyl-4-ethylthiazole substitutions shown. The SMILES code transformation: CC(C)n1cccn1>>CCc1nc(C)sc1C.

SMIRKS:
CC(C)n1cccn1>>Cc1sc(C)nc1CC

Thiopyrophosphoantigens: Solid-phase Synthesis and in Vitro Characterization of a New Class of Vγ9 Vδ2 T Cells Activators.[29]

In the process of developing new phosphoantigens, a number of bioisosteric compounds were considered. In the process, 1-bromo-2-methyl-2-butanol and 2-methyl-

but-2-ene were used. There is also consideration of a thiol group added or substituted in these compounds that could add to improved T-cell activation.

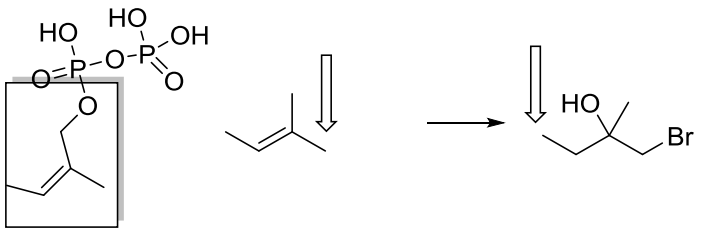

Figure 43. A prototype phosphoantigen with the substitution site shown. The substitution is 2-methyl-but-2-ene to 1-bromo-2-methyl-2-butanol. The SMILES code transformation: OCC(=C/C)/C>>BrCC(O)(C)CC

SMIRKS:
O[C:1][C:2](=[C:3][C:4])[C:5]>> Br[C:1][C:2](O)([C:3])[C:4][C:5]

Structural Evolutions of Salicylaldoximes as Selective Agonists for Estrogen Receptor β.[30]

A group of estrogen receptor ligands was synthesized by creating diaryl-substituted salicylaldoximes and anthranylaldoximes. Replacement of the α-napthol ring with a hydrogen-bonded pseudocyclic ring, combined with the addition of a chlorine and oxygen, led to strong affinity and selectivity for estrogen receptor β over estrogen receptor α.

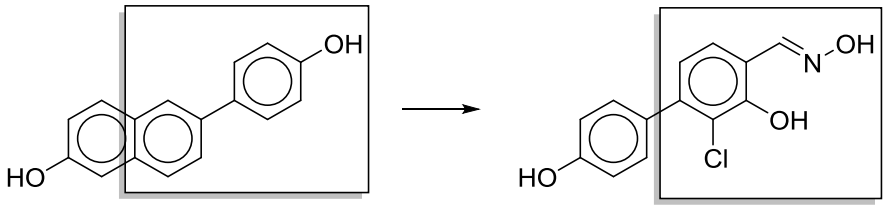

Figure 44. 6-(4-hydroxyphenyl)-2-naphthol to 2-chloro-4-[(E)-(hydroxyimino)methyl]-3,4'-biphenyldiol. The SMILES code transformation: Oc3ccc(c1ccc2c(c1)ccc(O)c2)cc3>>c1cc(ccc1c2ccc(c(c2Cl)O)/C=N/O)O

SMIRKS:
O[c:1]3[c:2]cc(c1ccc2c(c1)ccc(O)c2)cc3>>[c:1]1[c:2]c(ccc1c2ccc(c(c2Cl)O)/C=N/O)O

Analysis of the Reaction of Carbachol with Acetylcholinesterase Using Thioflavin T as a Coupled Fluorescence Reporter.[31]

This research focused on the reaction of carbachol (carbamoylcholine) and acetylcholinesterase (AChE), which hydrolyzes the neurotransmitter acetylcholine. In determining distinctive electronic differences between affinities of compounds, carbachol was compared with acetylcholine and acetylthiocholine as an isosteric analogue. Replacement of a methyl group with an amine is an isosteric replacement. In this case, the ester functional group becomes a carbamate, which adds stability to the molecule by reducing its propensity for hydrolysis.

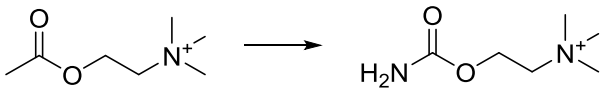

Figure 45. Acetylcholine to carbachol. The SMILES code transformation: CC(=O)OCC[N+](C)(C)C>>O=C(OCC[N+](C)(C)C)N.

SMIRKS:
C[C:1](=O)OCC[N+](C)(C)C>>O=C(OC[C:1][N+](C)(C)C)N

Pyrazolone-based anaplastic lymphoma kinase (ALK) inhibitors: Control of selectivity by a benzyloxy group.[32]

In the process of cancer screening, a number of Anaplastic Lymphoma Kinase (ALK) inhibitors were studied. Transmembrane receptor tyrosine kinase has been found to be active in several different cancers. ALK was compared to VEGFR2-based kinase. A loss of entropy in the bound form of the VEGFR-2 kinase in comparison with the former could explain the activity reduction. The addition of two dihydro-fluoro-difluoro groups also could play a significant role in this change in activity.

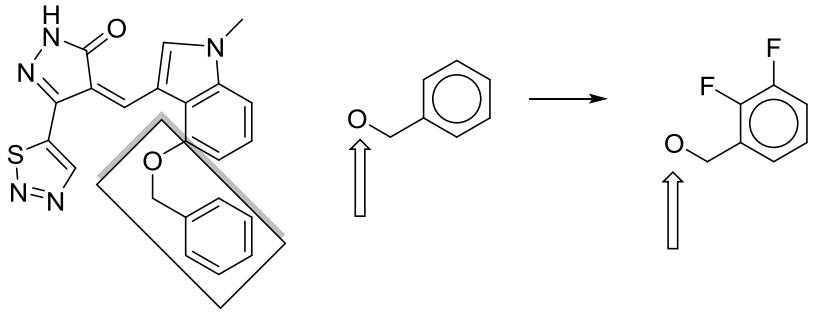

Figure 46. ALK inhibitor with substitution site shown, followed by benzyl alcohol to (2,3-difluorophenyl)methanol substitution. The hydrogen atoms are not shown in the figures, since the oxygens would not be protonated when connected to the main structure. The SMILES code transformation: c1ccccc(CO)1>>c1ccc(F)c(F)c(CO)1.

SMIRKS:
[c:1]1[c:2]=[c:3][c:4]=[c:5][c:6]([C:7][O:8])1>>[c:1]1[c:2]=[c:3](F)[c:4]=[c:5](F)[c:6]([C:7][O:8])1

Functionalized benzophenone, thiophene, pyridine, and fluorene thiosemicarbazone derivatives as inhibitors of cathepsin L.[33]

Cathepsin L is a protein that has been identified in a number of myofibral and myocardial processes. The isomeric substitutions shown below were present in active inhibitors of Cathepsin L without affecting Cathepsin B (a protein that has actually been identified as helpful in the battle against Alzheimers Disease). It is not unreasonable to assume that a 1,2-bromo shift would lead to other activity in other circumstances given the flexible nature of receptors.

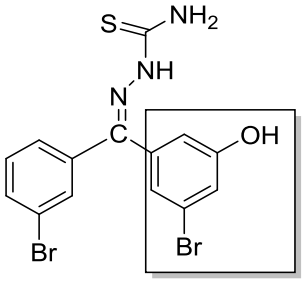

Figure 47. Benzophenone thiosemicarbazone cathepsin L inhibitor showing substitution site.

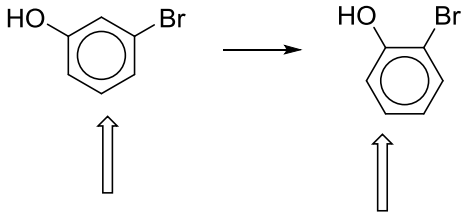

Figure 48. 3-bromophenol to 2-bromophenol substitution.
The SMILES code transformation: Oc1cc(Br)ccc1>>Oc1c(Br)cccc1

SMIRKS:
O[c:1]1[c:2][c:3](Br)[c:4][c:5][c:6]1>> O[c:1]1[c:2](Br)[c:3][c:4][c:5][c:6]1

DNA-dependent protein kinase (DNA-PK) inhibitors. Synthesis and biological activity of quinolin-4-one and pyridopyrimidin-4-one surrogates for the chromen-4-one chemotype.[34]

Dibenzo[b,d]thiophen-4-yl)-2-morpholino-4H-chromen-4-one (NU7441) was previously developed as a DNA-dependent protein kinase inhibitor. Using a variety of substitutions (shown below) on a scaffolding, the inhibition effects were observed in a number of the compounds. In the case of the two substituents (from Figure 50), they would be replacing the morpholine structure on the far right of Figure 49. An additional rule substitution using that molecule could be forthcoming in future coding.

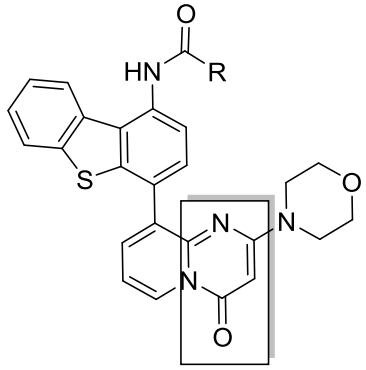

Figure 49. NU 7441, an antitumor agent and DNA-Protein Kinase inhibitor with site of substitution indicated.

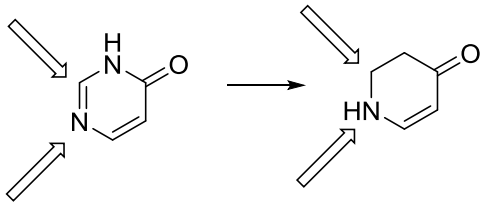

Figure 50. Compound pyrimidin-4(3H)-one with 2,3-dihydropyridin-4(1H)-one (1:1) substitution. The SMILES code transformation: O=C1C=CN=CN1>>O=C1C=CNCC1.

SMIRKS:
O=C1C=CN=CN1>>O=C1C=CNCC1

Exploring alternative Zn-binding groups in the design of HDAC inhibitors: Squaric acid, *N*-hydroxyurea, and oxazoline analogues of SAHA.[35]

Histone deacetylase inhibitors are used in a number of neurological and psychiatric drugs. A series of suberoylanilide hydroxamic acid (SAHA) substituted compounds was synthesized and evaluated for their inhibitory and cytotoxic activity. The compounds shown in Figure 52 would be substituted for the far right end of the molecule in Figure 51. Essentially, a five membered ring structure is replacing the hydroxyurea portion.

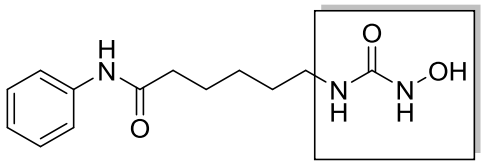

Figure 51. Substituted Suberoylanilide Hydroxamic Acid (SAHA) compound with transformation area highlighted.

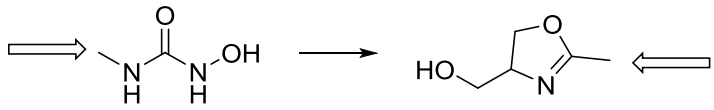

Figure 52. 1-Hydroxy-3- methylurea to (2-methyl-4,5-dihydro-1,3-oxazol-4-yl)methanol substitution. The SMILES code transformation:
CC1=NC(CO)CO1>>CNC(=O)NO

SMIRKS:
CC1=NC(CO)CO1>>ON(=O)NC

Design and evaluation of new antipsoriatic antedrug candidates having 16-en-22-oxa-vitamin D_3 structures.[36]

In an effort to treat psoriasis with minimal side effects, a group of substituted antidrugs were created. They increased antiproliferation activity, while lessening calcemic activity (in comparison to currently prescribed drugs). In these trials, a number of different compounds were attached to the lone carbon at the top of the scaffolding shown in Figure 53. The two figures shown in Figure 54 were among the most promising. The second compound in Figure 54 is similar in size and shape, but has significant differences in electronegativity due to the addition of five fluorine atoms.

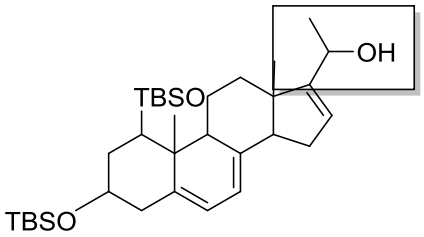

Figure 53. 16-en-22-oxa-vitamin D_3 scaffolding with substitution site. The substituted molecules replace the hydrogen atom shown in the hydroxide in the box. The TBSO on top of the six membered ring to the left of the figure is attached at the top of that ring, and not at the connection point between the two rings.

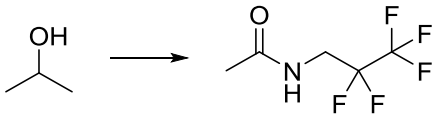

Figure 54. Propan-2-ol to N-(2,2,3,3,3-pentafluoropropyl)acetamide substitution. These were two substitutions (in addition to the replaced group shown in Figure 53). The SMILES code transformation:CC(O)C>>FC(F)(F)C(F)(F)CNC(=O)C

SMIRKS:
CC(O)(CC)CC>> FC(F)(F)C(F)(F)CNC(=O)C

Mechanistic Studies of Choline Oxidase with Betaine Aldehyde and Its Isosteric Analogue 3,3-Dimethylbutyraldehyde[37]

Studies involving choline inhibition have compared betaine aldehyde and 3,3-dimethylbutanal. The two compounds resemble each other sterically at the quaternary centers, with only the substitution in the center of the molecule being the difference between the two. Despite the geometric similarities, the 3,3-dimethylbutanal displayed significantly reduced flavin reduction in experiments performed by the researchers.

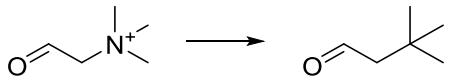

Figure 55. Betaine aldehyde to 3,3-dimethylbutyraldehyde substitution. The SMILES code transformation: O=CC[N+](C)(C)C>>O=CCC(C)(C)C.

SMIRKS:
[O:1]=[C:2][C:3][N+]([C:4])([C:5])([C:6])>>[O:1]=[C:2][C:3]C([C:4])([C:5])([C:6])

Reactions of Oxetan-3-*tert*-butylsulfinimine for the Preparation of Substituted 3-Aminooxetanes.[38]

An oxetane ring has been used in a number of bioisosteric substitutions involving dimethyl and carbonyl groups. In this study, 3-aminooxetanes were developed through reactions involving oxetan-3-*tert*-butylsulfinimine and the corresponding sulfinylaziridine. In the preparation of these molecules, the compounds in Figure 56 were attached to sulfinimines to create 3-aminooxetanes that could serve as bioisosteric replacements for geminal dimethyl groups. The larger phenyl group was substituted in this case, and may have limited use, but the transformation was noted.

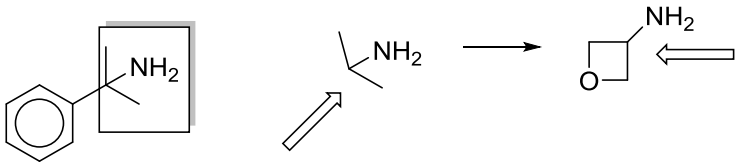

Figure 56. Parent structure with 2-propanamine to 3-oxetanamine with attachment points. The SMILES code transformation: CC(N)(C)>>NC1COC1.

SMIRKS:
CC([N:1])(C)[c:2]1ccccc1>>[N:1][C:2]1COC1

Synthesis of a 6-Methyl-7-deaza Analogue of Adenosine that Potently Inhibits Replication of Polio and Dengue Viruses.[39]

Tubercidin is a natural glycolysis inhibitor. Mimics of adenosine and adenosine triphosphate (ATP) have been found to be effective in therapeutics. Synthesis of a tubercidin analogue was found to be an acceptable substitution for adenosine in a number of categories. 6-Methyl-9-β-D-ribofuranosylpurine was eventually isolated and also found to be an effective substitution. The substitution of a carbon atom for a nitrogen atom in the aromatic ring was the only change.

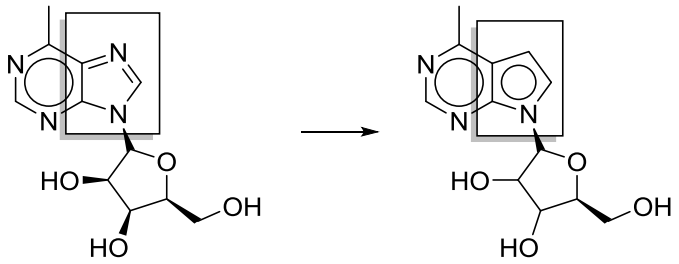

Figure 57. Adenosine to tubercidin (7-(β-D-Lyxofuranosyl)-7H-pyrrolo[2,3-d]pyrimidin-4-amine). The SMILES code transformation:
c1nc(c2c(n1)N(cn2)[C@H]3[C@@H]([C@@H]([C@H](O3)CO)O)O)C >>Cc3ncnc2c3ccn2[C@@H]1O[C@H](CO)C(O)C1O

SMIRKS:
c1nc(c2c(n1)n(cn2)[CH]3[C]([C]([C](O3)CO)O)O)N>>Nc3ncnc2c3ccn2[C]1O[C](CO)C(O)C1O

The Steric Hypothesis for DNA Replication and Fluorine Hydrogen Bonding Revisited in Light of Structural Data.[40]

Researchers studying DNA replication have hypothesized that molecular geometry is more important than hydrogen bonding in certain situations. Using 2,4–difluorotoluene as a bioisosteric substitute for thymine seemed to verify the idea because fluorine was thought to be a poor hydrogen bond acceptor. In this study, however, it was determined that both the geometry and hydrogen bonding are equally important in the substitutions, and the fluoro group is capable of engaging in hydrogen bonding in many cases.

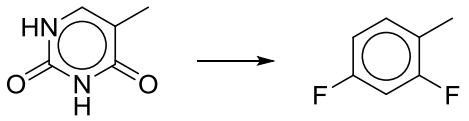

Figure 58. Thymine to 2,4 – difluorotoluene substitution. The SMILES code transformation: Cc1c[nH]c(=O)[nH]c1=O>>Fc1cc(F)c(cc1)C.

SMIRKS:
C[c:1]1c[nH]c(=O)[nH]c1=O>>Fc1cc(F)c(c[c:1]1)C

Thermodynamic Insights on the Structure and Energetics of *s*-Triphenyltriazine.[41]

When testing the two compounds shown below, researchers found that they are isosteres. The triazine, however, has a higher stability because of the geometry. The authors suggested that the geometry decreased the volatility, which could explain the increase in stability. This led to the conclusion that ring-ring torsion effects played a major portion in the decrease in stability for the latter. Aside from the change in geometry, there is also a significant substitution in the center ring, where three nitrogen atoms have been replaced with carbon atoms.

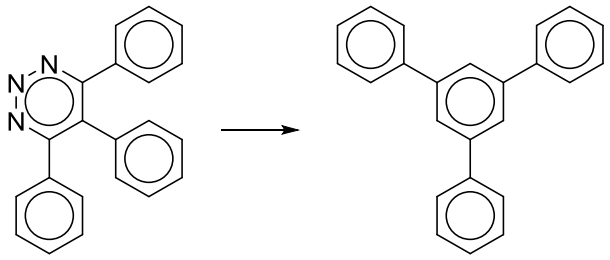

Figure 59. 2,4,6-Triphenyl-1,3,5-triazine to 1,3,5-triphenylbenzene substitution. The SMILES code transformation:
n2nnc(c1ccccc1)c(c2c3ccccc3)c4ccccc4>>c1cc(ccc1)c2cc(cc(c2)c3ccccc3)c4ccccc4

SMIRKS:
n2nnc([c:1]1[c:2][c:3][c:4][c:5][c:6]1)c(c2c3ccccc3)c4ccccc4>>[c:1]1[c:2][c:3]([c:4][c:5][c:6]1)c2cc(cc(c2)c3ccccc3)c4ccccc4

Synthesis of new 4-aminoquinolines and quinoline–acridine hybrids as antimalarial agents.[42]

Chloroquine (CQ) is one of the two most widely used antimalarial drugs in the world. It is a relatively inexpensive drug and is readily available. Unfortunately, resistance has been steadily increasing towards CQ. Substitutions were made to 4-aminoquinolines and quinolone-acridine (acridine shown in Figure 60 and chloroquine shown in Figure 61) hybrids. The substituted compounds were found to be equal to four times as powerful as CQ in dosage. The transformation in Figure 63 involves little in terms of changing elements, but involves a bringing together of two ring structures, which will have a difference in steric effects.

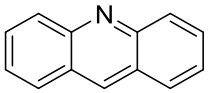

Figure 60. Example of an acridine molecule.

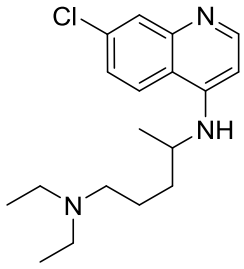

Figure 61. Chloroquine is used to treat malaria.

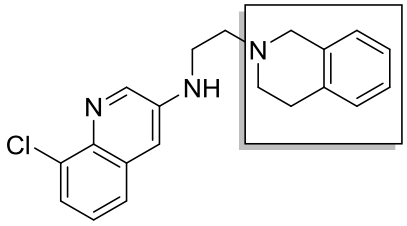

Figure 62. Molecule with site of substitution.

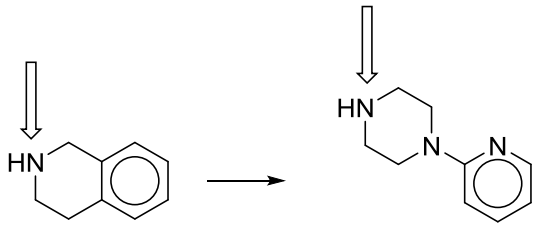

Figure 63. Compound 1,2,3,4-tetrahydroisoquinoline to 1-(2-pyridinyl)piperazine. The arrows indicate the points of attachment. The SMILES code transformation: C1Cc2ccccc2CN1>>c1nc(ccc1)N2CCNCC2.

SMIRKS:
[C:1]1Cc2ccccc2CN1>>[c:1]1nc(ccc1)N2CCNCC2

4-Piperidines and 3-pyrrolidines as dual serotonin and noradrenaline reuptake inhibitors: design, synthesis and structure-activity relationships.[43]

Structure-activity relationships with piperidine and pyrrolidine derivatives as norepinephrine reuptake inhibitors (NRI's) and serotonin norepinephrine reuptake inhibitors (SNRI's) were investigated. It was found that dual ring substitution allowed for selection of NRI's and SNRI's. Additional pharmacological studies are ongoing. As can be seen in Figure 64 below, the nitrogen has changed positions with a carbon in the

aromatic ring structure, and the six member piperidine ring has been replaced with a five member pyrrolidine ring.

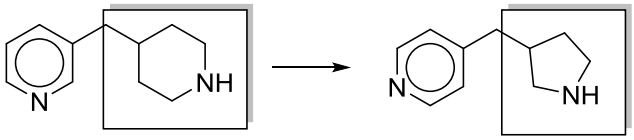

Figure 64. Compound 3-(4-Piperidinylmethyl)pyridine to 4-(3-pyrrolidinylmethyl)pyridine substitution. The SMILES code transformation: c1ncccc1CC2CCNCC2>>C(C1CCNC1)c2ccncc2.

SMIRKS:
c1[n:1][c:2][c:3]cc1CC2CCNCC2>>C(C1CCNC1)c2cc[n:1][c:2][c:3]2

Substituted *N*-{3-[(1,1-dioxido-1,2-benzothiazol-3-yl)(phenyl)amino]propyl}benzamide analogs as potent Kv1.3 ion channel blockers. Part 2.[44]

Kv1.3 is a potassium voltage-gated channel that is important in a number of body regulations, including heart rate, secretion of insulin, muscle contraction, and neurotransmitter release. With such a variety of functions, the ability to regulate this channel is essential. Researchers are developing potential Kv1.3 channel blockers using *N*-3-[(1,1-dioxido-1,2-benzothiazol-3-yl)(phenyl)amino]propylbenzamide analogs. Kv1.3 channel blockers using *N*-3-[(1,1-dioxido-1,2-benzothiazol-3-yl)(phenyl)amino]propyl-benzamide analogs. The substitutions are shown are at the far right end of the molecule in Figure 65

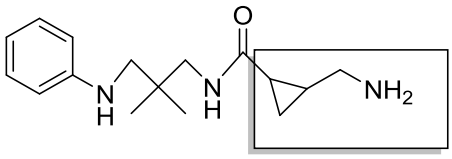

Figure 65. Site of substitution for 1-Cyclopropylmethanamine to 1,1,1-Trifluoro-2-propanol.

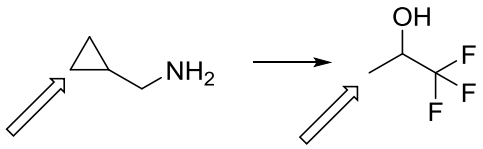

Figure 66. 1-Cyclopropylmethanamine to 1,1,1-Trifluoro-2-propanol with attachment points shown by arrows. The SMILES code transformation: NCC1CC1>>CC(O)C(F)(F)F.

SMIRKS:
NCC1CC1>>CC(O)C(F)(F)F

Replacement of pyrazol-3-yl amine hinge binder with thiazol-2-yl amine: Discovery of potent and selective JAK2 inhibitors.[45]

The human protein, janus kinase 2 (commonly known as JAK2) is believed to be involved in growth regulation. The lack of this protein has been shown to kill laboratory mice within 12 days. Regulation of this protein is very important. When working with a number of potential inhibitors, thiazol-2-yl amine was selected as a potential isosteric replacement for pyrazol-3-yl amine. The structures are very similar, but the arrangements of the atoms on the ring structure (and the sulfur substitution for nitrogen) give the molecules different activities.

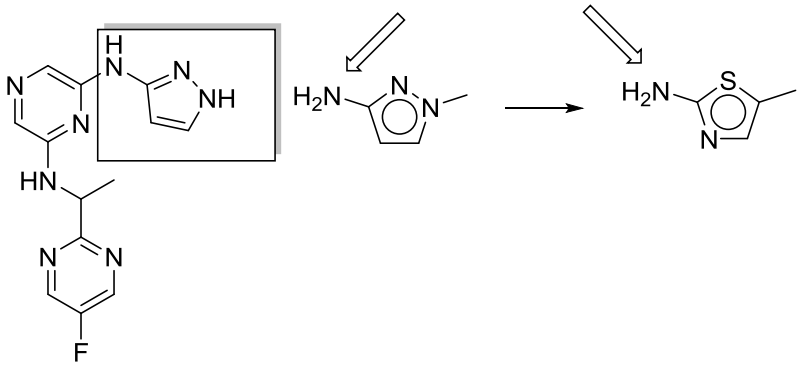

Figure 67. Parent structure with 1-methyl-1H-pyrazol-3-amine to 5-methyl-1,3-thiazol-2-amine substitution. The SMILES code transformation: Nc1ccn(C)n1>>Cc1cnc(N)s1.

SMIRKS:
[N:1][c:2]1ccn(C)n1>> Cc1cn[c:2]([N:1])s

Structural Insights into the Dual Activities of the Nerve Agent Degrading Organophosphate Anhydrolase/Prolidase.[46]

The bimetalloenzyme organophosphate acid anhydrolase (OPAA) is known to inhibit a number of acetylcholinesterase-inhibiting organophosphorus compounds, including toxic nerve agents. In an effort to obtain a better understanding of how this compound works, glycolate tendencies were observed and noted. The glycolate compounds appear to be very active and may explain some of OPAA's tendencies. The substitution is minor (the hydroxyl group replaces the amine to the far left of the molecule) yet it seems to affect the activity considerably.

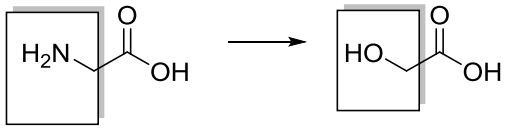

Figure 68. Glycine to glycolic acid transformation. The SMILES code transformation: C(C(=O)O)N>>C(C(=O)O)O.

SMIRKS:
[C:1]([C:2](=[O:3])[O:4])N>>[C:1]([C:2](=[O:3])[[O:4]])O

Novel substituted (*Z*)-5-((*N*-benzyl-1*H*-indol-3-yl)methylene)imidazolidine-2,4-diones and 5-((*N*-benzyl-1*H*-indol-3-yl)methylene)pyrimidine-2,4,6(1*H*,3*H*,5*H*)-triones as potent radio-sensitizing agents.[47]

HT-29 is an aggressive adenocarcinoma found in colorectal cancers. A series of derivatives was analyzed for their ability to reduce activity of the HT-29 cell line. The substituted portions of these derivatives were the cyclic compounds shown below. The substitutions for the groups include a larger ring structure in the compound to the right (see Figure 69) and an additional oxygen atom double bonded to the ring.

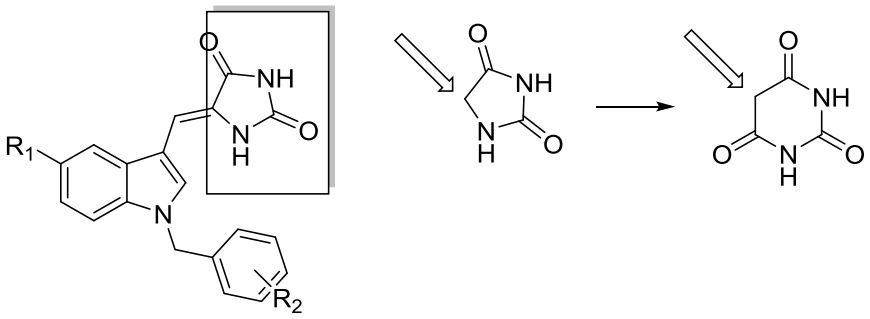

Figure 69. Parent compound with substitution area (boxed) with imidazolidine-2,4-dione to pyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione transformation, with attachment points indicated by arrows. The SMILES code formatting:
O=C1NC(=O)CN1>>O=C1CC(=O)NC(=O)N1.

SMIRKS:
O=C1NC(=O)CN1>> O=C1CC(=O)NC(=O)N1

Pyrazolopyridazine alpha-2-delta-1 ligands for the treatment of neuropathic pain[48]

Researchers used a variety of substituted pyrazolopyridazine alpha-2-delta-1 ligands in an effort to control neuropathic pain. In the case of the substitution shown in Figure 70, a chlorine atom is substituted with a fluorine atom on the left side of the aromatic ring. The trifluoromethyl group on the right of the structure has been replaced with a two carbon chain. While the size of these molecules are very similar, their activities are quite different.

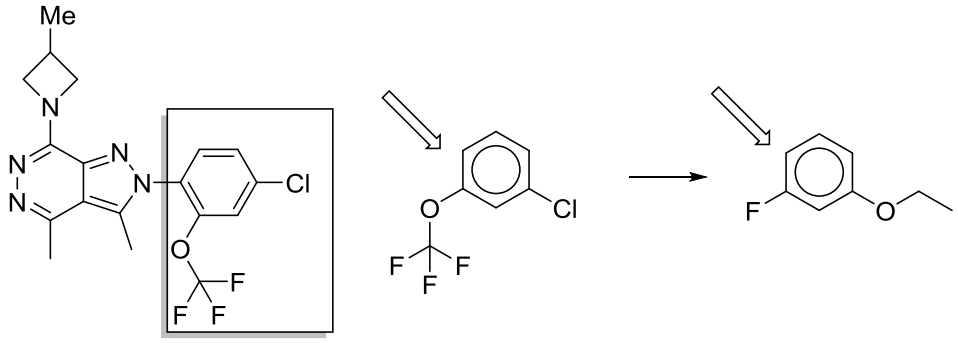

Figure 70. Parent structure with 3-chlorophenyl trifluoromethyl ether to 1-ethoxy-3-fluorobenzene substitution. The SMILES code transformation:
Clc1cc(OC(F)(F)F)ccc1>>CCOc1cccc(F)c1.

SMIRKS:
Clc1cc(OC(F)(F)F)ccc1>> CCOc1cccc(F)c1

Recognition Properties of Carboxylic Acid Bioisosteres: Anion Binding by Tetrazoles, Aryl Sulfonamides, and Acyl Sulfonamides on a Calix[4]arene Scaffold.[49]

Tetrazoles, aryl sulfonamides, and acyl sulfonamides are often used as substitutions for carboxylic acids in drug development. These compounds often show improved stability and potency. Presumably, the acyl group when bonded to the sulfonamides makes the N-H group more acidic (i.e., decreases the pKa), which enhances the solubility and mimics a carboxylate group. The substitution in Figure 71 is coming off of the aromatic ring, where a tetrazole is being replaced by an acyl sulfonamide.

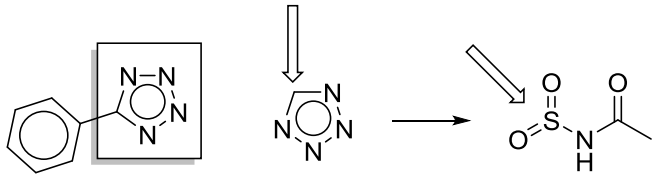

Figure 71. 5-phenyl-1H-tetrazole with substituted compounds. The hydrogen attached to the sulfur has been removed, since it will not be present when attached to the parent structure. The SMILES code transformation: n1nnnc1>>O=S(=O)(NC(=O)C)

SMIRKS:
n1nnnc1>>O=S(=O)(NC(=O)C)1

2-Aryl-4,5,6,7-tetrahydro-1,3-benzothiazol-7-ols as a class of antitumor agents selectively active in securin$^{-/-}$ cells.[50]

A group of 2-(4-aminophenyl)-4,5,6,7-tetrahydro-1,3-benzothiazol-7-ol derivatives was investigated as potential antitumor agents. A substitution was made in the six membered ring with the replacement of a carbon atom with a nitrogen atom. That substitution seemed to have a definite impact on the antitumor abilities.

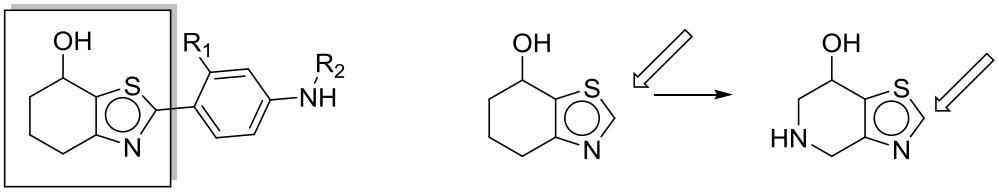

Figure 72. Parent structure with 4,5,6,7-tetrahydro-1,3-benzothiazol-7-ol to 4,5,6,7-tetrahydro[1,3]thiazolo[4,5-c]pyridin-7-ol substitution. The SMILES code transformation: OC1CCCc2ncsc12>>OC1CNCc2ncsc12.

SMIRKS:
O[C:1]1[C:2]C[C:3][c:4]2[n:5][c:6][s:7][c:8]12>>O[C:1]1[C:2]N[C:3][c:4]2[n:5][c:6][s:7][c:8]12

Design and synthesis of cell potent BACE-1 inhibitors: structure-activity relationship of P1' substituents.[51]

A number of hydroxyethylamine BACE-1 inhibitors were manipulated to increase cathepsin-D inhibition. The compounds examined were six membered rings with a variety of attachments and compounds. The most effective substitution is shown in the figure below.

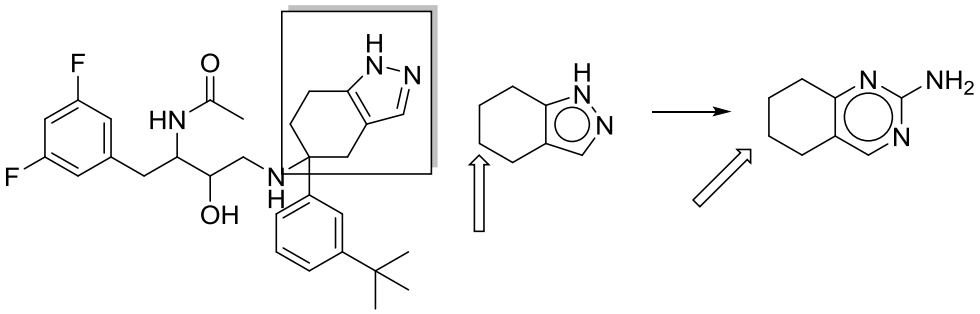

Figure 73. Parent structure with 4,5,6,7-Tetrahydro-1H-indazole to 5,6,7,8-Tetrahydro-2-quinazolinamine substitution. The SMILES code transformation: C1CCc2[nH]ncc2C1>>Nc1nc2CCCCc2cn1

SMIRKS:
[C:1]1CCc2[nH]ncc2C1>>Nc1nc2[C:1]CCCc2cn1

Chiral N^G-acylated hetarylpropylguanidine-type histamine H_2 receptor agonists do not show significant stereoselectivity.[52]

A group of imidazolylpropylguanidines and 2-aminothiazolylpropylguanidines were substituted and synthesized in an effort to produce potent histamine H_2 receptor (H_2R) agonists. The substituted compound had a similar size to the original compound (with only a sulfur atom replacing a nitrogen atom), but there are also a nitrogen and carbon atom attached to the aromatic ring that may have increased the agonist's potency.

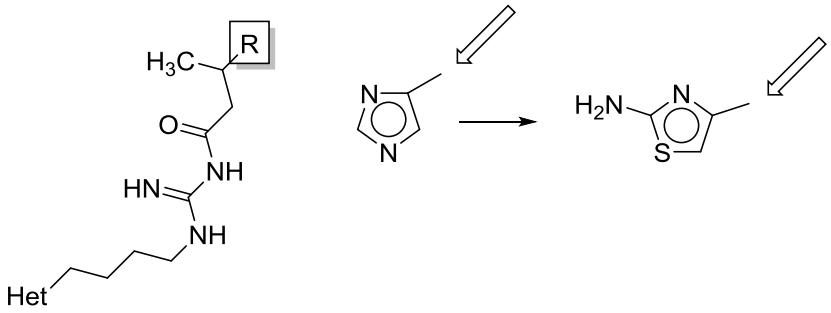

Figure 74. Parent structure with 1*H*-imidazole to aminomethiazole substitution. The SMILES code transformation: c1cn(C)cn1>>Cc1csc(N)n1.

SMIRKS:
[c:1]1[c:2][n:3](C)cn1>>C[c:1]1[c:2]sc(N)[n:3]1

1,2,3-Triazole–Heme Interactions in Cytochrome P450: Functionally Competent Triazole–Water–Heme Complexes.[53]

1,2,4-Triazole is a recognized Cytochrome P450 inhibitor. For whatever reason, 1,2,3-triazole has not received the same status. A number of pharmaceuticals utilize 1,2,4-triazole for heme-nitrogen interactions, but none use the 1,2,3-triazole. When substituted onto a scaffold, it was found that the ligand positioning was altered. The 1,2,3-triazole did not exhibit the same inhibitory effects, but it did offer a unique

regioselectivity that could allow it to be used in other situations. The transformation here is simply a rearrangement of nitrogen and carbon placement on the triazole's ring structure.

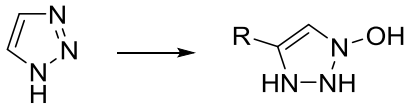

Figure 75. 1,2,3-triazole to triazole substitution. The SMILES code transformation: N1=NNC=C1>>ON1=NNC(*)=C1.

SMIRKS:
n1[c:1][c:2][n:3][n:4]1>> c1[c:1][n:2][n:3][nH]1

Design and Synthesis of Celecoxib and Rofecoxib Analogues as Selective Cyclooxygenase-2 (COX-2) Inhibitors: Replacement of Sulfonamide and Methylsulfonyl Pharmacophores by an Azido Bioisostere.[54]

In an effort to design selective cyclooxygenase-2 (COX-2) Inhibitors, researchers modified celecoxib and rofecoxib analogues by using the bioisosteric substitution shown below. The substitution increased anti-inflammatory and analgesic activities in the analogues tested. The substituted compound had a similar five-member cyclic ring system, but had a number of atom changes (including the removal of all fluorine and nitrogen atoms, some of which were replaced by oxygen atoms).

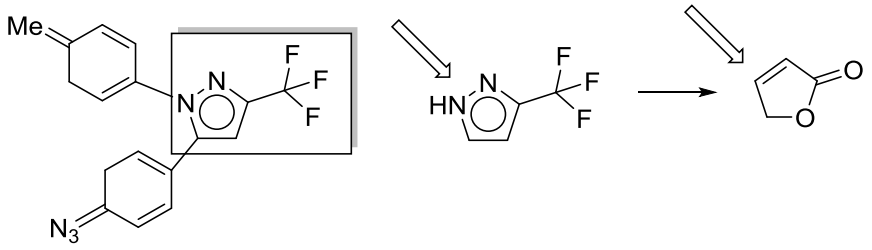

Figure 76. Parent structure with 5-(trifluoromethyl)-1H-pyrazole to 2(5H)-furanone substitution. The SMILES code transformation: c1c[nH]nc1C(F)(F)F>>O=C\1OC/C=C/1.

SMIRKS:
[c:1]1c[nH]nc1C(F)(F)F>>O=[C:1]\1OC/C=C/1

Exploring Aromatic Chemical Space with NEAT: Novel and Electronically Equivalent Aromatic Template.[55]

Computer modeling simulations were used to develop biosisosteres. Properties such as electrostatic potential charges, reactivity, hydrogen bonding ability, and dipole moments were explored when calculating potential substituent ring structures. The transformation of 2-methyl-1-benzofuran to 2-methyl-1,3a-dihydropyrazolo[1,5-*a*]pyridine was found to fit these qualifications. The ring structures are very similar in size, but have different abilities based upon the substitution of nitrogen atoms for the oxygen and carbon atoms in the original structure and the change in bonding. The areas circled indicate the points of attachment.

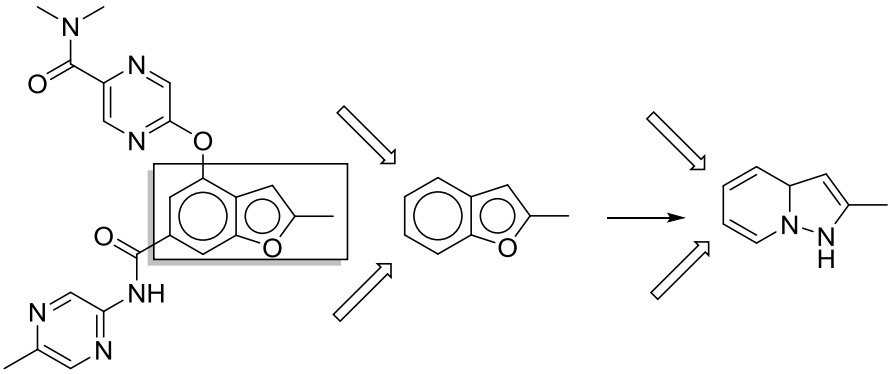

Figure 77. Parent structure with 2-methyl-1-benzofuran to 2-methyl-1,3a-dihydropyrazolo[1,5-*a*]pyridine substitution. The SMILES code transformation: Cc1cc2ccccc2o1>>CC1=CC2C=CC=CN2N1.

SMIRKS:
[C:1][c:2]1[c:3]c2ccccc2o1>>[C:1][C:2]1=[C:3]C2C=CC=CN2N1

Identification of target family directed bioisosteric replacements.[56]

Forty protein families were analyzed as potential bioisosteres. These bioisosteres were isolated and found to target a dozen different groups including nuclear hormone receptors and tyrosine kinases. In this transformation, the sulfur atom was replaced by a carbon atom, and the five-member ring was expanded to a six member ring.

Figure 78. 2-R-Thiophene to 3-*R*-1,4-cyclohexadiene substitution. The SMILES code transformation: s1cccc(*)1>>c1ccccc(*)1.

SMIRKS:
s1[c:1][c:2][c:3][c:4](*)1>>[c:1]1[c:2][c:3][c:4]cc(*)1

***In vitro* and *in vivo* antimalarial evaluation of semi-synthetic derivatives of gomphostenin.[57]**

Sathe and coworkers are actively studying gomphostenins (compounds that are derived from the gomphostemma plant) for their antimalarial properties. By making substitutions, two of these derivatives were found to actually have more success than gomphostenins themselves. The substitution was one substituent switch of a fluorine atom for a chlorine atom. This should not affect the overall size or shape of the molecule significantly, but the greater electronegativity of a fluoro group would alter the electrostatics.

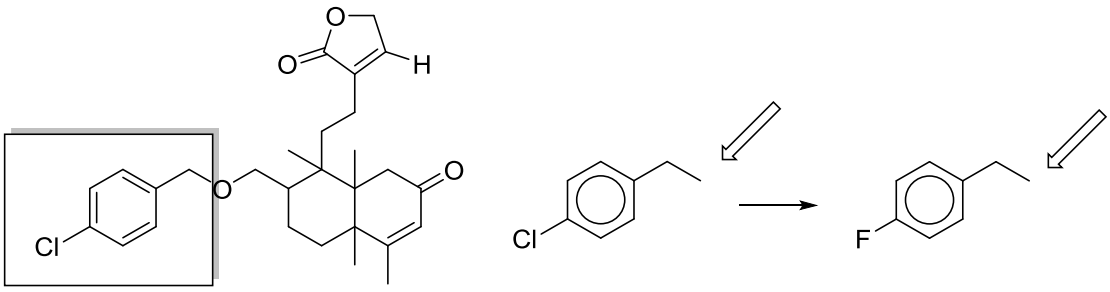

Figure 79. Parent structure with 1-Chloro-4-ethylbenzene to 1-Fluoro-4-ethylbenzene substitution. The SMILES code transformation: CCc1ccc(Cl)cc1>>CCc1ccc(F)cc1.

SMIRKS:
[C:1][C:2][c:3]1[c:4][c:5][c:6](Cl)[c:7][c:8]1>>[C:1][C:2][c:3]1[c:4][c:5][c:6](F)[c:7][c:8]1

Substituted isoxazole analogs of farnesoid X receptor (FXR) agonist GW4064.[58]

Starting from the known FXR (farnesoid X receptor, a sensor in the bile that works with nuclear receptors in bile acid uptake, metabolism and excretion regulation) agonist GW 4064 1a, a series of alternately 3,5-substituted isoxazoles was prepared as FXR agonists. The substitutions involved the cyclocompounds shown below. With either substituted ring compound, the potency remained the same. For this transformation, the aromatic ring remains constant, but the attachments completely changed. The two chlorine atoms were replaced by two carbon atoms and a hydroxyl group.

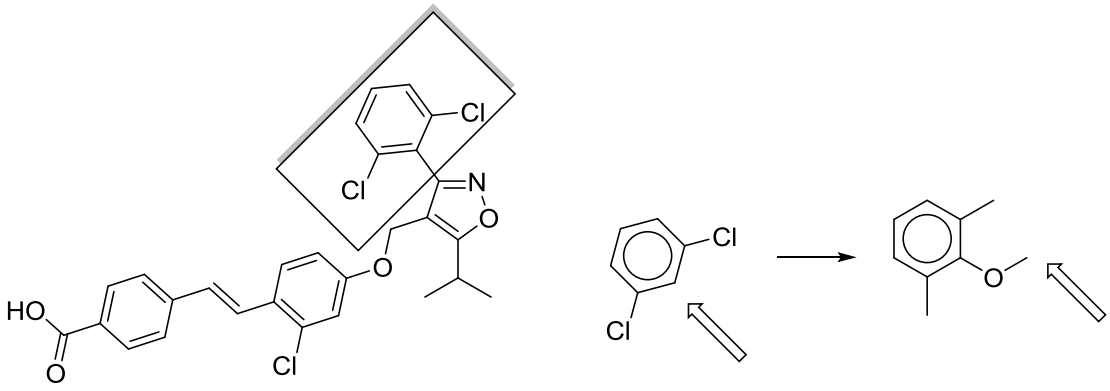

Figure 80. Parent compound GW 4064 with the substitution site indicated followed by the transformation of 1,3-dichlorobenzene to 2-methoxy-1,3-dimethylbenzene.
The SMILES code transformation: Clc1cccc(Cl)c1>>Cc1c(OC)c(C)ccc1.

SMIRKS:
Cl[c:1]1[c:2][c:3][c:4][c:5](Cl)[c:6]1>>C[c:1]1[c:2](OC)[c:3](C)[c:4][c:5][c:6]1

Tricyclic sulfones as orally active γ-secretase inhibitors: Synthesis and structure–activity relationship studies.[59]

In a design study of γ-secretase inhibitors, a series of substituted tricyclic sulfones (including BMS-708163, as shown below) was developed and found to be very potent. This potency could be used in the future development of new pharmaceuticals for the treatment of Alzheimer's disease. In the transformation, the chlorine atom was substituted for the trifluoromethyl group.

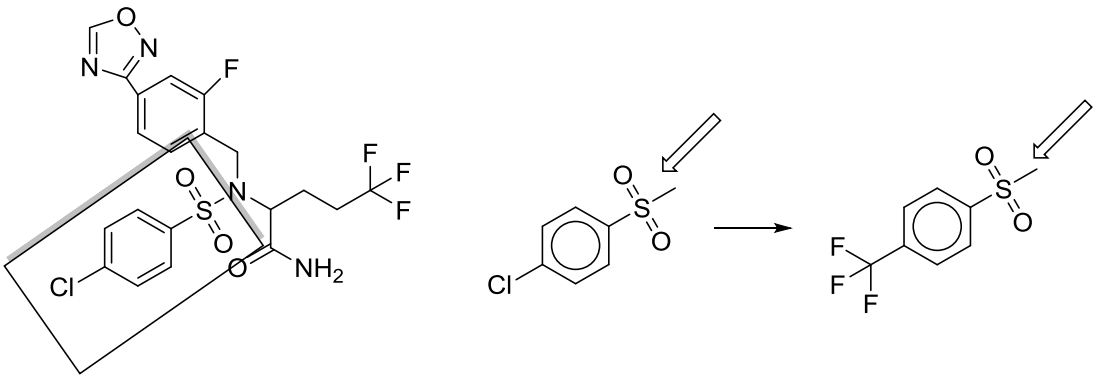

Figure 81. BMS 708163 (with substitution site indicated) followed by transformation of 1-chloro-4-(methylsulfonyl)benzene to 1-(methylsulfonyl)-4-(trifluoromethyl)benzene. The SMILES code transformation:
O=S(C)(=O)c1ccc(Cl)cc1>>O=S(C)(=O)c1ccc(cc1)C(F)(F)F

SMIRKS:
[O:1]=[S:2]([C:3])(=[O:4])[c:5]1[c:6][c:7][c:8](Cl)([c:9][c:10]1)
>>[O:1]=[S:2]([C:3])(=[O:4])[c:5]1[c:6][c:7][c:8]([c:9][c:10]1)C(F)(F)F

Pyridylmethylthio derivatives as VEGF inhibitors. Part 1[60]

Vascular endothelial growth factor (VEGF) is important in vasculogenesis and angiogenesis. Normally, it is useful in creating new blood vessels. Unfortunately, when overexpressed, it can lead to a number of vascular diseases. It can also provide blood flow to cancers that would otherwise be unable to survive. In this study, a number of

pyridylmethylthio substituted compounds were developed and found to be very active in VEGF inhibition. In this case, the aromatic ring remains constant, but a trifluoromethyoxy group has replaced two carbon attachments.

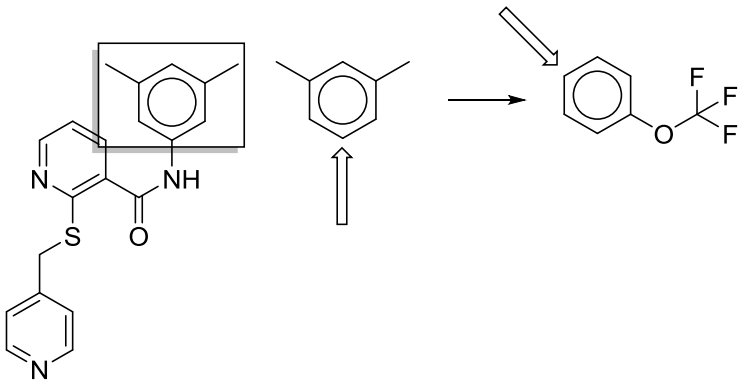

Figure 82. Parent inhibitor with substitution site indicated (box) followed by m-xylene to (trifluoromethoxy)benzene transformation. The SMILES code transformation: Cc1cccc(C)c1>>FC(F)(F)Oc1ccccc1.

SMIRKS:
C[c:1]1[c:2][c:3][c:4][c:5](C)[c:6]1>>FC(F)(F)O[c:1]1[c:2][c:3][c:4][c:5][c:6]1

Application of the Bicyclo[1.1.1]pentane Motif as a Nonclassical Phenyl Ring Bioisostere in the Design of a Potent and Orally Active γ-Secretase Inhibitor.[61]

While working with γ-secretase inhibitor (BMS-708163), researchers decided to substitute Bicyclo[1.1.1]pentane for the fluorobenzene that was currently in the compound. When the substitution was made, it led to a compound that had the same potency as before, but with increases in passive permeability and aqueous solubility. This shows that the bicycle ring has capapilities beyond just being a “spacer”.

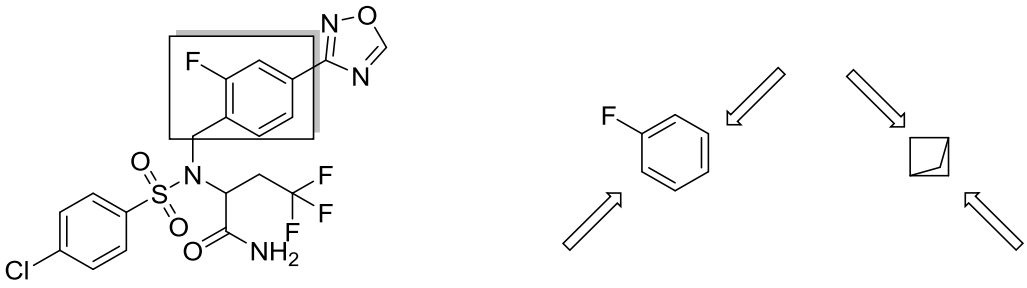

Figure 83. BMS 708163 with substitution site indicated followed by a phenyl ring (in this case, fluorobenzene) to bicyclo[1.1.1]pentane. The arrows indicate attachment points. The SMILES code transformation: Fc1ccccc1>>C1C2CC1C2.

SMIRKS:
Fc1ccccc1>> C1C2CC1C2

Novel thienopyrrole glycogen phosphorylase inhibitors: Synthesis, in vitro SAR and crystallographic studies.[62]

In an effort to design a potent glycogen phosphorylase a (GPa) inhibitor, a number of thienopyrrole substituted compounds were synthesized. These substitutions led to reduced glucose output from rat hepatocytes. For the transformation, the aromatic six member ring structure remains constant, but the attached ring expands from a five member carbon ring to a six member ring with a nitrogen substituent and an attached oxygen.

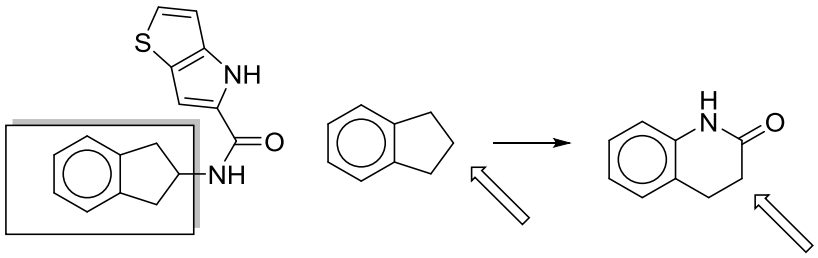

Figure 84. Inhibitor with substitution point indicated followed by transformation of indane to 3,4-dihydro-2(1H)-quinolinone. The SMILES code transformation: c1cccc2CCCc12>>O=C1CCc2ccccc2N1.

SMIRKS:
[c:1]1[c:2][c:3][c:4][c:5]2CCC[c:6]12>>O=C1CC[c:1]2[c:2][c:3][c:4][c:5][c:6]2N1

Discovery of novel 1-phenyl-cycloalkane carbamides as potent and selective influenza fusion inhibitors.[63]

Hemagglutinin (HA) is a protein that binds to sialic acid receptors on cells and is active in the influenza virus. Researchers used substituted 1-phenyl-cycloalkane carbamides to develop a potent HA inhibitor. The transformation from the piperidine keeps the ring structure, but converts to an aromatic structure with the addition of two side chains. The circles in Figure 86 indicate the points of attachment.

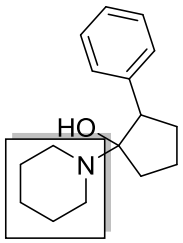

Figure 85. Structure showing the piperidine molecule's location before substitution.

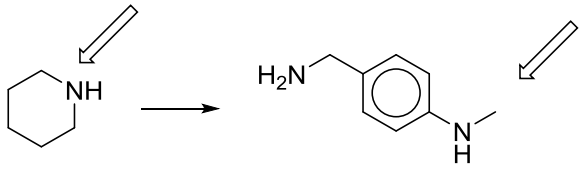

Figure 86. Piperidine to 4-(aminomethyl)-*N*-methylaniline transformation. The SMILES code transformation: C1CCCCN1>>c1c(NC)ccc(CN)c1.

SMIRKS:
C1CCCCN1>> CNc1ccc(CN)cc1

Ureas with histamine H_3-antagonist receptor activity—A new scaffold discovered by lead-hopping from cinnamic acid amides.[64]

Histamine H_3 (hH_3) receptors are active in the central and peripheral nervous systems as neurotransmitter modulators. Treatment of these receptors has opened up a wide range of potential drugs for pharmacologic intervention for a wide range of diseases, ranging from obesity to sleep apnea. In this study, a series of substituted urea compounds was utilized as hH_3 antagonists. The substitution in this transformation is on the left side of the molecules. The carbon-carbon double bond is replaced by a carbon atom attached to a nitrogen atom.

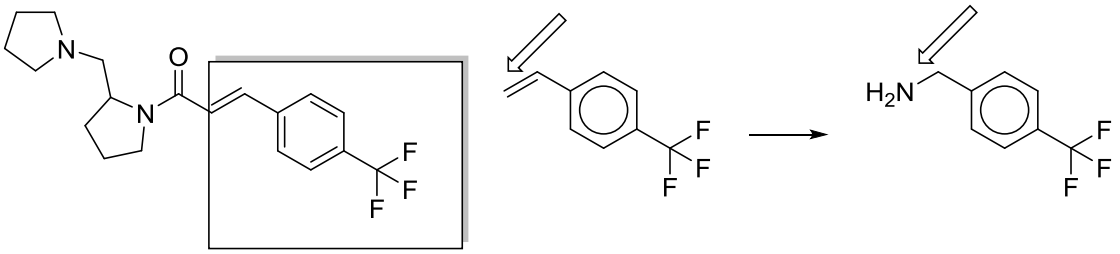

Figure 87. Parent compound with substitution site (box) followed by 4-vinylbenzotrifluoride to 1-[4-(trifluoromethyl)phenyl]methanamine transformation. The SMILES code transformation: FC(F)(F)c1ccc(C=C)cc1>>FC(F)(F)c1ccc(cc1)CN.

SMIRKS:
[F:1][C:2]([F:3])([F:4])[c:5]1[c:6][c:7][c:8](C=C)cc1>>
[F:1][C:2]([F:3])([F:4])[c:5]1[c:6][c:7][c:8](cc1)CN

Cyclopentane-1,3-dione: A Novel Isostere for the Carboxylic Acid Functional Group. Application to the Design of Potent Thromboxane (A2) Receptor Antagonists.[65]

Because cyclopentane-1,3-dione has a pK_a value similar to a carboxylic acid, researchers investigated the possibility of using it as a bioisostere. When comparing the properties of the compound to other carboxylic acid isosteres (such as tetrazole), it matched up favorably. A number of experiments were carried out to determine if cyclopentane-1,3-dione would be a suitable isostere. With its acidity and versatility, it could certainly be a strong candidate for future isosteric replacements.

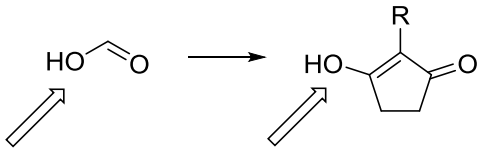

Figure 88. Carboxylic acid to cyclopentane-1,3-dione substitution. The SMILES code transformation: OC=O>>O=C1CCC(O)=C1(*)

SMIRKS:
*C(=O)O>> O=C1CCC=(=O)C1

Oxamate is an Alternative Substrate for Pyruvate Carboxylase from *Rhizobium etli*.[66]

Oxamate is actively involved in the enzymatic decarboxylation of oxaloacetate in the pyruvate carboxylase domain. It has been shown to be isosterically and isoelectronically similar to pyruvate. The replacement of a methyl with a primary amino group has been shown to be an effective isosteric replacement. In this case, the remainder of the molecule is the same, aside from that change. Based upon the similar structures, masses, and sizes, these two compounds should be interchangeable in a number of situations.

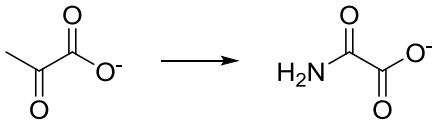

Figure 89. Pyruvate to oxamate substitution. The SMILES code transformation: CC(=O)C(=O)[O-]>>[O-]C(=O)C(=O)N.

SMIRKS:
C[C:1](=O)[C:2](=O)[O-]>>[O-][C:1](=O)[C:2](=O)N

Dihydroorotase from the Hyperthermophile *Aquifiex aeolicus* is Activated by Stoichiometric Association with Aspartate Transcarbamoylase and Forms a One-Pot Reactor for Pyrimidine Biosynthesis.[67]

In mammals, pyrimidine biosynthesis is vital in cellular metabolism. Carbamoyl phosphate synthetase is one of the first three enzymes in this production. Citrate becomes attached in the course of this process. It is a near isosteric analogue of carbamoyl aspartate. In the transformation shown in Figure 90, the molecule's shape has changed slightly, but the nitrogen atoms have been completely removed, with one hydroxyl group being added.

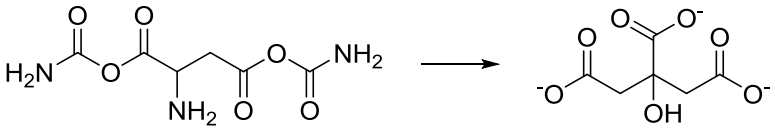

Figure 90. Dicarbamoyl aspartate to citrate substitution. The SMILES code transformation: NC(CC(=O)OC(N)=O)C(=O)OC(N)=O>>C(C(=O)[O-])C(CC(=O)[O-])(C(=O)[O-])O

SMIRKS:
NC([C:1][C:2](=O)OC(N)=O)C(=O)OC(N)=O>>C(C(=O)[O-])C([C:1][C:2](=O)[O-])(C(=O)[O-])O

***N*-Hydroxypyrazolyl Glycine Derivatives as Selective *N*-Methyl-D-aspartic Acid Receptor Ligands.[68]**

A number of analogues were designed and produced utilizing *N*-hydroxypyrazole as a functional bioisostere for the distal carboxylate group on aspartate. Because so many

excitatory signals in the brain are influenced by glutamic acid through the glutamate receptors, there are many associated neurological problems. These include epilepsy, schizophrenia, and Alzheimer's disease. With bioisosteric substitution, it is thought that improved selectivity can be achieved in neurotransmission. The substitution shown below could lead to improved selectivity in activation of NMDA receptors. In the transformation, an aromatic ring structure (pyrazol) has replaced the two oxygen atoms on the far right of the figure (See Figure 91).

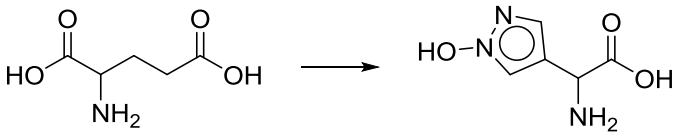

Figure 91. Glutamic acid to amino(1-hydroxy-1H-pyrazol-4-yl)acetic acid substitution. The SMILES code transformation:
C(CC(=O)O)C(C(=O)O)N>>c1c(cn(n1)O)C(C(=O)O)N.

SMIRKS:
C(CC(=O)O)[C]([C:1](=O)O)N>>N[C:1](C(=O)O)c1cn(O)nc1

Exploration of novel thiobarbituric acid-, rhodanine- and thiohydantoin-based HIV-1 integrase inhibitors.[69]

Virtual screening techniques were used to identify compounds capable of inhibiting HIV-1 integrase. The molecules shown below were attached to larger scaffoldings. These molecules were the portions that were actively substituted. In the transformation shown below, a six member closed ring was reduced to a five member

ring, a sulfur atom was substituted for a nitrogen atom, and an attached oxygen was removed.

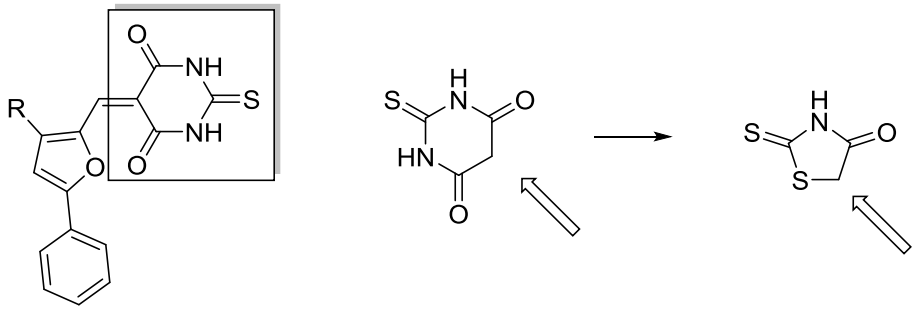

Figure 92. Thiobarbituric acid to rhodanine substitution. The parent compound is to the left. The SMILES code transformation:
O=C1CC(=O)NC(=S)N1>>O=C1CSC(=S)N1.

SMIRKS:
[O:1]=[C:2]1CC(=O)NC(=S)N1>>[O:1]=[C:2]1CSC(=S)N1

A Diverse Series of Substituted Benzenesulfonamides as Aldose Reductase Inhibitors with Antioxidant Activity: Design, Synthesis, and in Vitro Activity.[70]

A phenosulfonamido acetic acid chemotype was determined to be an effective aldose reductase inhibitor. Replacement of the acetic acid with a difluorophenol exhibited potent antioxidant behavior. With the replacement of a small chain with a six member aromatic compound, selectivity could still an issue, so further adjustments need to be made. This could be a substitution that is limited by size issues and could only take place where plenty of room is available.

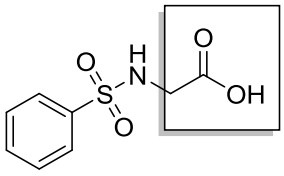

Figure 93. Benzenesulfonamide with substituted group indicated.

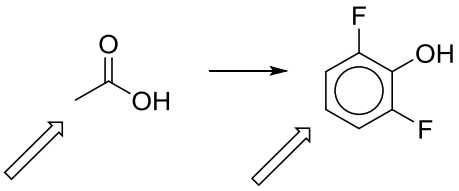

Figure 94. Acetic acid to 2,6-difluorophenol, with the attachment points shown with arrows. The SMILES code transformation: CC(=O)O>>c1(F)c(O)c(F)ccc1.

SMIRKS:
CC(=O)[O:1]>> c1(F)c([O:1])c(F)ccc1

Ursodeoxycholic Acid Amides As Novel Glucocorticoid Receptor Modulators [71]

Ursodeoxycholic acid (UDCA) treats inflammatory diseases. Researchers tested UDCA derivatives to examine their effects on cancer lines. For this transformation, a hydroxyl group was replaced by an amide with a three member closed ring.

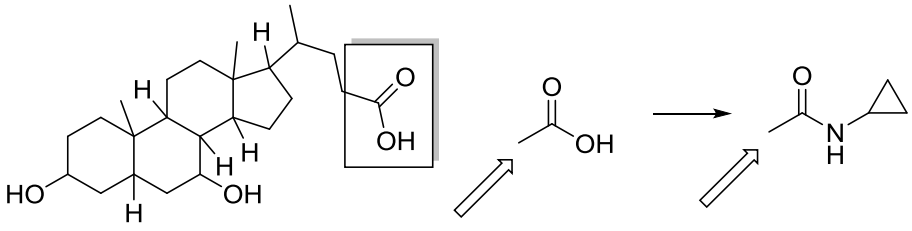

Figure 95. UDCA with the substitution site indicated, followed by the transformation of acetic acid to *N*-cyclopropylacetamide. The SMILES code transformation: CC(=O)O>>CC(=O)NC1CC1.

SMIRKS:
CC(=O)O>>CC(=O)NC1CC1

Separation of Quercetin's Biological Activity from Its Oxidative Property through Bioisosteric Replacement of the Catecholic Hydroxyl Groups with Fluorine Atoms.[72]

Quercetin has been shown to have significant biologic activity due to its strong antioxidative properties. It is an abundant flavonoid, found in an abundance of foods. Oxidative damage is often associated with melanomas. The researchers wanted to isolate the various parts of quercetin by bioisosteric substitution. By doing this, they hope to figure out what specific part of the compound is responsible for the strong antioxidative state. A number of substitutions were tried in various parts of this large structure. In one successful bioisosteric substitution, two fluorine atoms replaced two hydroxides, giving the compound slightly more biological activity.

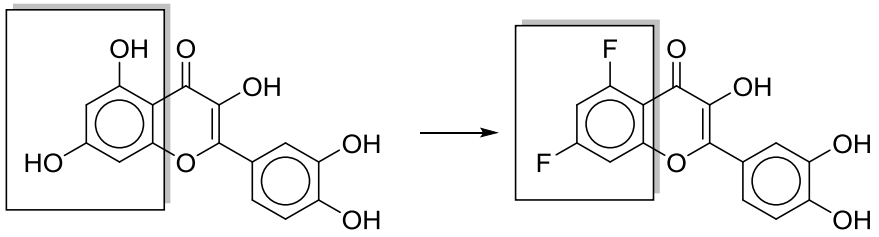

Figure 96. Quercetin to 3,4-difluoro-quercetin substitution. The SMILES code transformation:
O=C1c3c(O/C(=C1/O)c2ccc(O)c(O)c2)cc(O)cc3O>>O=C1c3c(O/C(=C1/O)c2ccc(O)c(O)c2)cc(F)cc3F.

SMIRKS:
[O:1]=[C:2]1[c:3]3[c:4](O/C(=C1/O)c2ccc(O)c(O)c2)cc(O)cc3O>>
[O:1]=[C:2]1[c:3]3[c:4](O/C(=C1/O)c2ccc(O)c(O)c2)cc(F)cc3F

Inhibitors for Human Glutaminyl Cyclase by Structure Based Design and Bioisosteric Replacement.[73]

Inhibition of human glutamyl cyclase (hQC) has been identified as a potential treatment for Alzheimer's disease. A transformation of 1-(3,4-Dimethoxyphenyl)-3-[3-(1H-imidazol-1-yl)propyl]thiourea to N-(3,4-dimethoxyphenyl)-1-[3-(1H-imidazol-1-yl)propyl]cyclopropanecarbothioamide is shown to significantly increase inhibition, by a simple addition of a three carbon ring.

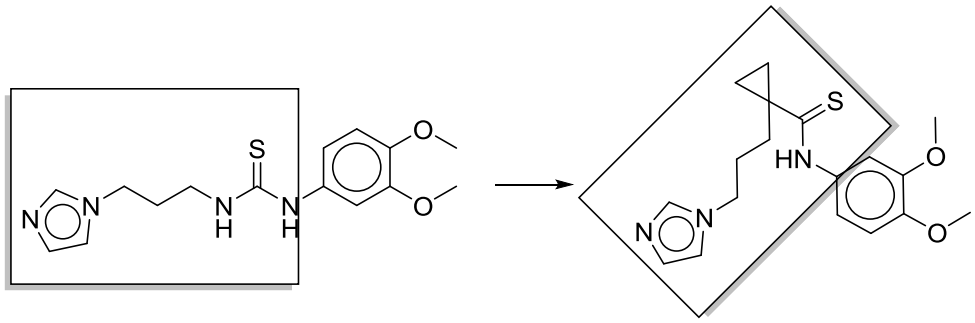

Figure 97. 1-(3,4-Dimethoxyphenyl)-3-[3-(1H-imidazol-1-yl)propyl]thiourea to *N*-(3,4-dimethoxyphenyl)-1-[3-(1*H*-imidazol-1-yl)propyl]cyclopropanecarbothioamide substitution. The SMILES code transformation: S=C(NCCCn1ccnc1)Nc2ccc(OC)c(OC)c2>>COc1ccc(cc1OC)NC(=S)C3(CCCn2ccnc2)CC3.

SMIRKS: [S:1]=[C:2](NCCCn1ccnc1)Nc2ccc(OC)c(OC)c2>>COc1ccc(cc1OC)N[C:2](=[S:1])C3(CCCn2ccnc2)CC3

CHAPTER VII

CONCLUSION

In the growing field of cheminformatics, utilizing computer software to facilitate pharmaceutical drug design is essential. The SMILES/SMIRKS coding (used in the Drug Guru program) is user friendly and offers many opportunities for additional software development. As more research efforts are directed toward the advancements of this drug design approach, the software will improve, the coding will become even simpler, and the ability of the software to eliminate unreasonable choices and to rank the transformations will make the system even more effective.

There were a number of new transformations presented in the previous pages. A number of the transformations are good for selective situations, while another quantity is more open to a larger scale. While some of the transformations may seem very selective, in fact they will become more usable as other connections are made. Consider two compounds that are related to each other structurally; when a third compound is found to have bioisosteric similarities to a compound, it may also become attached to the compound that was connected to it. These sorts of connections will keep the database growing constantly. A specific example is the transformation shown in Figure 29 earlier in the paper. This transformation could be called a carbonyl ene shift. In other words, when you have a conjugated enone system, the positions of the carbonyl and the double bond could be switched. This would take only a simple modification of the code

The transformations presented in this thesis were numerous, but certainly not comprehensive. The Bowen research group in the Center for Drug Design at Mercer University College of Pharmacy will continue to investigate the possibilities for improvement in cheminformatics software and development. As more transformations are entered into the group's database, the potential for bioisosteric substituents increases. This means that there could be theoretical an almost endless supply of bioisosteric transformations that could find their way into the database. So, while this is a conclusion to this thesis, it is only a small building block into the construction of a much larger future for drug design.

REFERENCES

1. Kadam, S.S.; Mahadik, K.R.; Bothara, K.G. *Principles of Medicinal Chemistry, Volume 2.* Gultekdi, India: Nirali Prakashan, 2007. 403. Print.
2. Lemke, Thomas; Williams, David A.; Roche, Victoria F.; Zito, S. William; *Foye's Principles of Medicinal Chemistry, Sixth Edition.* Baltimore: Lippincott, 2008: 513. Print.
3. Kar, Ashutash; *Medicinal Chemistry, Revised Third Edition.* New Delhi, India: New Age International Publishers. 2005. 45-47. Print.
4. Nogrady, Thomas; Weaver, Donald F.; *Medicinal Chemistry: A Molecular and Biochemical Approach, Third Edition.* New York: Oxford University Press. 2005. 21. Print.
5. Regis, Ed; *The Info Mesa: Science, Business, and New Age Alchemy on the Santa Fe Plateau.* New York: W.W. Norton & Company. 2003. 82-91. Print.
6. Stewart, Kent D.; Shiroda, Melisa; James, Craig A.; "Drug Guru: A computer software program for drug design using medicinal chemistry rules." *Bioorganic and Medicinal Chemistry* 14 (2006): 7011-7022.
7. Brown, Nathan; *Bioisosteres in Medicinal Chemistry.* Weinheim, Germany. Wiley VCH, Verlag & Company. 2012. 10-15. Digital.
8. Judson, Philip; *Knowledge-based Expert Systems In Chemistry: Not Counting On Computers.* Cambridge, England: Royal Society of Chemistry. 2009. 45. Print.
9. Silverman, Richard B.; *The Organic Chemistry of Drug Design and Drug Action.* Oxford, England. Elsevier Press. 2004. Print.
10. Adler, Marc; Baldwin, Steven. "Direct, regioselective synthesis of 2,2-dimethyl-2H-chromenes. Total syntheses of octandrenolone and precocenes I and II." *Tetrahedron Letters* 50 (2009) 5075–5079.
11. Juli, Christina, et al. "Pipecolic Acid Derivatives As Small-Molecule Inhibitors of the *Legionella* MIP Protein." *Journal of Medicinal Chemistry* (2011): 277-283.
12. Ochoa-Puentes, Cristian, et al. "Benzanilide–Biphenyl Replacement: A Bioisosteric Approach to Quinoline Carboxamide-Type ABCG2 Modulators." *ACS Medicinal Chemistry Letters* (2013): 393-396.
13. Matta, Cherif F., Arabi, Alya A., Weaver, Donald F.; "The biosteric similarity of the tetrazole and carboxylate anions: Clues from the topologies of the electrostatic potential and of the electron density.

14. Gonzalez-Lafuente, Laura, et al. “Benzothiazepine CGP37157 and its isosteric 2'-methyl analogue provide neuroprotection and block cell calcium entry.” *ACS Chemical Neuroscience* (2012): 519-529.
15. Clarke, Brian, et al. “BACE-1 hydroxyethylamine inhibitors using novel edge-to-face interaction with Arg-296.” *Bioorganic & Medicinal Chemistry Letters* (2010): 4639-4644.
16. Gundugola, Aditya S., et al. “Synthesis and antiproliferative evaluation of 5-*oxo* and 5-thio derivatives of 1,4-diaryl tetrazoles.” *Bioorganic & Medicinal Chemistry Letters* (2010): 3920-3924.
17. Kumar, Dalip, et al. “Synthesis and biological evaluation of indolyl chalcones as antitumor agents.” *Bioorganic & Medicinal Chemistry Letters* (2010): 3916-3919.
18. Le, Giang, et al. “Design of a series of bicyclic HIV-1 integrase inhibitors. Part 2: Azoles: Effective metal chelators.” *Bioorganic and Medicinal Chemistry Letters* (2010): 6909-6912.
19. Pruksakom, Patamapom, et al. “Trichoderins, novel aminolipopeptides from a marine sponge-derived *Trichoderma* sp., are active against dormant mycobacteria.” *Bioorganic & Medicinal Chemistry Letters* (2010): 3658-3663.
20. Khalaj, Ali, et al. “Synthesis, aerobic cytotoxicity, and radiosensitizing activity of novel 2,4-dinitrophenylamine tethered 5-fluorouracil and hydroxyurea.” *Bioorganic & Medicinal Chemistry Letters* (2006): 6034-6038.
21. Burgey, Christopher S., et al. “Benzodiazepine calcitonin gene-related peptide (CGRP) receptor antagonists: Optimization of the 4-substituted piperidine.” *Bioorganic & Medicinal Chemistry Letters* (2006): 5052-5056.
22. Hall, Matthew D., et al. “Synthesis, activity, and pharmacophore development for isatin-beta-thiosemicarbazones with selective activity toward multidrug-resistant cells.” *Journal of Medicinal Chemistry* (2009): 3191-3204.
23. Sinning, Christian, et al. “New Analgesics Synthetically Derived from the Paracetamol Metabolite *N*-(4-Hydroxyphenyl)-(5*Z*,8*Z*,11*Z*,14*Z*)-icosatetra-5,8,11,14-enamide.” *Journal of Medicinal Chemistry* (2008): 7800-7805.
24. Reiner, Andreas, et al. “Effect of Thioxopeptide Bonds on α-Helix Structure and Stability.” *Journal of the American Chemical Society* (2008): 8079-8084.
25. Yang, Hanbiao, et al. “Cyclic amide bioisosterism: Strategic application to the design and synthesis of HCV NS5B polymerase inhibitors.” *Bioorganic and Medicinal Chemistry Letters* (2010): 4614-4619.

26. Chauhan, Shikha S., et al. “Synthesis and biological evaluation of indolyl glyoxylamides as a new class of antileishmanial agents.” *Bioorganic & Medicinal Chemistry Letters* (2010): 6191-6194.
27. Therrien, Eric, et al. “1,2-Diamines as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1).” *Bioorganic & Medicinal Chemistry Letters* (2009): 6725-6732.
28. Goble, Stephen D., et al. “Heterocyclic acetamide and benzamide derivatives as potent and selective β_3-adrenergic receptor agonists with improved rodent pharmacokinetic profiles.” *Bioorganic and Medicinal Chemistry Letters* (2010): 1895-1899.
29. Breccia, Perla, et al. “Thiopyrophosphoantigens: Solid-phase Synthesis and in Vitro Characterization of a New Class of Vγ9 Vδ2 T Cells Activators.” *Journal of Medicinal Chemistry* (2009): 3716-3722.
30. Minutolo, Filippo, et al. “Structural Evolutions of Salicylaldoximes as Selective Agonists for Estrogen Receptor β.” *Journal of Medicinal Chemistry* (2009): 858-867.
31. Rosenberry, Terrone L., et al. “Analysis of the Reaction of Carbachol with Acetylcholinesterase Using Thioflavin T As a Coupled Fluorescence Reporter.” *Biochemistry* (2008): 13056-13063.
32. Tripathy, Rabindranath, et al. “Pyrazolone-based anaplastic lymphoma kinase (ALK) inhibitors: Control of selectivity by a benzyloxy group.” *Bioorganic & Medicinal Chemistry Letters* (2011): 7261-7264.
33. Kumar, G.D. Kishore, et al. “Functionalized benzophenone, thiophene, pyridine, and fluorene thiosemicarbazone derivatives as inhibitors of cathepsin L.” *Bioorganic & Medicinal Chemistry Letters* (2010): 6610-6615.
34. Cano, Celine, et al. “DNA-dependent protein kinase (DNA-PK) inhibitors. Synthesis and biological activity of quinolin-4-one and pyridopyrimidin-4-one surrogates for the chromen-4-one chemotype.” *Journal of Medicinal Chemistry* (2010): 8498-8507.
35. Hanessian, Stephan, et al. “Exploring alternative Zn-binding groups in the design of HDAC inhibitors: Squaric acid, *N*-hydroxyurea, and oxazoline analogues of SAHA.” *Bioorganic & Medicinal Chemistry Letters* (2006): 4784-4787.
36. Shimizu, Kasuki, et al. “Design and evaluation of new antipsoriatic antedrug candidates having 16-en-22-oxa-vitamin D_3 structures.” *Bioorganic & Medicinal Chemistry Letters* (2006): 3323-3329.
37. Hamzik, Philip J.; Brubaker, Jason D. “Reactions of Oxetan-3-*tert*-butylsulfinimine for the Preparation of Substituted 3-Aminooxetanes.” *Organic Letters* (2010): 1116-1119.

38. Fan, Fan; Germann, Markus W.; Gadda, Giovanni. “Mechanistic Studies of Choline Oxidase with Betaine Aldehyde and Its Isosteric Analogue 3,3-Dimethylbutyraldehyde” *Biochemistry* (2006): 1979-1986.
39. Wu, Runzhi, et al. “Synthesis of a 6-Methyl-7-deaza Analogue of Adenosine That Potently Inhibits Replication of Polio and Dengue Viruses.” *Journal of Medicinal Chemistry* (2010): 7958-7966.
40. Egli, Martin. “The Steric Hypothesis for DNA Replication and Fluorine Hydrogen Bonding Revisited in Light of Structural Data.” *Accounts of Chemical Research* (2012): 1237-1246.
41. Lima, Carlos F. R. A. C., Costa, Jose C. S., Santos, Luis M. N. B. F. “Thermodynamic Insights on the Structure and Energetics of *s*-Triphenyltriazine.” *The Journal of Physical Chemistry* (2011): 9249-9258.
42. Kumar, Ashok, et al. “Synthesis of new 4-aminoquinolines and quinoline–acridine hybrids as antimalarial agents.” *Bioorganic & Medicinal Chemistry Letters* (2010): 7059-7063.
43. Fish, Paul V. et al. “4-Piperidines and 3-pyrrolidines as dual serotonin and noradrenaline reuptake inhibitors: design, synthesis and structure-activity relationships.” *Bioorganic & Medicinal Chemistry Letters* (2009): 2829-2834.
44. Haffner, Curt D., et al. “Substituted *N*-{3-[(1,1-dioxido-1,2-benzothiazol-3-yl)(phenyl)amino]propyl}benzamide analogs as potent Kv1.3 ion channel blockers. Part 2.” *Bioorganic & Medicinal Chemistry Letters* (2010): 6989-6993.
45. Ionnidis, Stephanos, et al. “Replacement of pyrazol-3-yl amine hinge binder with thiazol-2-yl amine: Discovery of potent and selective JAK2 inhibitors.” *Bioorganic and Medicinal Chemistry Letters* (2010): 1669-1673.
46. Vyas, Nand K., et al. “Structural Insights into the Dual Activities of the Nerve Agent Degrading Organophosphate Anhydrolase/Prolidase.” *Biochemistry* (2010): 547-559.
47. Reddy, Y.T., et al. “Novel substituted (*Z*)-5-((*N*-benzyl-1*H*-indol-3-yl)methylene)imidazolidine-2,4-diones and 5-((*N*-benzyl-1*H*-indol-3-yl)methylene)pyrimidine-2,4,6(1*H*,3*H*,5*H*)-triones as potent radio-sensitizing agents.” *Bioorganic & Medicinal Chemistry Letters* (2010): 600-602.
48. Myatt, James W., et al. “Pyrazolopyridazine alpha-2-delta-1 ligands for the treatment of neuropathic pain.” *Bioorganic & Medicinal Chemistry Letters* (2010): 4683-4688.
49. Pinter, Thomas, et al. “Recognition Properties of Carboxylic Acid Bioisosteres: Anion Binding by Tetrazoles, Aryl Sulfonamides, and Acyl Sulfonamides on a Calix[4]arene Scaffold.” *The Journal of Organic Chemistry* (2011): 3733-3741.

50. Zhang, Nan, et al. “2-Aryl-4,5,6,7-tetrahydro-1,3-benzothiazol-7-ols as a class of antitumor agents selectively active in securin$^{-/-}$ cells.” *Bioorganic & Medicinal Chemistry Letters* (2010): 3903-3905.
51. Sealy, Jennifer M., et al. “Design and synthesis of cell potent BACE-1 inhibitors: structure-activity relationship of P1' substituents.” *Bioorganic & Medicinal Chemistry Letters* (2009): 6386-6391.
52. Ghorai, Prasanta, et al. “Chiral N^G-acylated hetarylpropylguanidine-type histamine H_2 receptor agonists do not show significant stereoselectivity.” *Bioorganic and Medicinal Chemistry Letters* (2010): 3173-3176.
53. Conner, Kip P., et al. “1,2,3-Triazole–Heme Interactions in Cytochrome P450: Functionally Competent Triazole–Water–Heme Complexes.” *Biochemistry* (2012): 6441-6457.
54. Habeeb, Amgad G., Praveen Rao, P. N., Knaus, Edward N. “Design and Synthesis of Celecoxib and Rofecoxib Analogues as Selective Cyclooxygenase-2 (COX-2) Inhibitors: Replacement of Sulfonamide and Methylsulfonyl Pharmacophores by an Azido Bioisostere.” *Journal of Medicinal Chemistry* (2001): 3039-3042.
55. Tu, Meihua, et al. “Exploring Aromatic Chemical Space with NEAT: Novel and Electronically Equivalent Aromatic Template.” *Journal of Chemical Information and Modeling* (2012): 1114-1123.
56. Wassermann, Anne Mai; Bajorath, Jürgen. “Identification of target family directed bioisosteric replacements.” *MedChemComm* 2 (2011): 601-606.
57. Sathe, Manisha; Kaushik, M. P. “In vitro and in vivo antimalarial evaluation of semi-synthetic derivatives of gomphostenin.” *Bioorganic & Medicinal Chemistry Letters* (2010): 4233-4236.
58. Bass, Jonathan Y., et al. “Substituted isoxazole analogs of farnesoid X receptor (FXR) agonist GW4064.” *Bioorganic & Medicinal Chemistry Letters* (2009): 3078-3080.
59. Sasikumar, T. K., et al. “Tricyclic sulfones as orally active γ-secretase inhibitors: Synthesis and structure–activity relationship studies.” *Bioorganic and Medicinal Chemistry Letters* (2010): 3632-3636.
60. Tajimi, Hisashi, et al. “Pyridylmethylthio derivatives as VEGF inhibitors. Part 1.” *Bioorganic & Medicinal Chemistry Letters* (2010): 7234-7238.
61. Stepan, Antonia F., et al. “Application of the Bicyclo[1.1.1]pentane Motif as a Nonclassical Phenyl Ring Bioisostere in the Design of a Potent and Orally Active γ-Secretase Inhibitor.” *Journal of Medicinal Chemistry* (2012): 3414-3424.

62. Whittamore, Paul R.O., et al. “Novel thienopyrrole glycogen phosphorylase inhibitors: Synthesis, in vitro SAR and crystallographic studies.” *Bioorganic & Medicinal Chemistry Letters* (2006): 5567-5571.
63. Tang, Guozhi, et al. “Discovery of novel 1-phenyl-cycloalkane carbamides as potent and selective influenza fusion inhibitors.” *Bioorganic and Medicinal Chemistry Letters* (2010): 3507-3510.
64. Lau, Jesper F., et al. “Ureas with histamine H_3-antagonist receptor activity—A new scaffold discovered by lead-hopping from cinnamic acid amides.” *Bioorganic & Medicinal Chemistry Letters* (2006): 5303-5308.
65. Ballatore, Carlo, et al. “Cyclopentane-1,3-dione: A Novel Isostere for the Carboxylic Acid Functional Group. Application to the Design of Potent Thromboxane (A2) Receptor Antagonists.” *Journal of Medicinal Chemistry* (2011): 6969-6983.
66. Marlier, John F., Cleland, W. W., Zeczycki, Tonya N. “Oxamate Is an Alternative Substrate for Pyruvate Carboxylase from *Rhizobium etli.”Biochemistry* (2013): 2888-2894.
67. Zhang, Pengfei, et al. “Dihydroorotase from the Hyperthermophile *Aquifiex aeolicus* Is Activated by Stoichiometric Association with Aspartate Transcarbamoylase and Forms a One-Pot Reactor for Pyrimidine Biosynthesis.” *Biochemistry* (2009): 766-778.
68. Clausen, Rasmus P., et al. “*N*-Hydroxypyrazolyl Glycine Derivatives as Selective *N*-Methyl-D-aspartic Acid Receptor Ligands.” *Journal of Medicinal Chemistry* (2008): 4179-4187.
69. Rajamake, Suvi, et al. “Exploration of novel thiobarbituric acid-, rhodanine- and thiohydantoin-based HIV-1 integrase inhibitors.” *Bioorganic & Medicinal Chemistry Letters* (2009): 3616-3618.
70. Alexiou, Polyxeni P., Demopoulos, Vassillis J. “A Diverse Series of Substituted Benzenesulfonamides as Aldose Reductase Inhibitors with Antioxidant Activity: Design, Synthesis, and in Vitro Activity.” *Journal of Medicinal Chemistry* (2010): 7756-7766.
71. Sharma, Ruchika, et al. “Ursodeoxycholic Acid Amides As Novel Glucocorticoid Receptor Modulators.” *Journal of Medicinal Chemistry* (2011): 122-130.
72. Cho, Suh Young, et al. “Separation of Quercetin’s Biological Activity from Its Oxidative Property through Bioisosteric Replacement of the Catecholic Hydroxyl Groups with Fluorine Atoms.” *Journal of Agricultural and Food Chemistry* (2012): 6499-6506.

73. Buchholz, Mirko, et al. “Inhibitors for Human Glutaminyl Cyclase by Structure Based Design and Bioisosteric Replacement.” *Journal of Medicinal Chemistry* (2009): 7069-7080.